1-1-2015

On the Adoption Dynamics of Internet Technologies: Models and Case Studies

Mehdi Nikkhah
University of Pennsylvania, mn5979@gmail.com

On the Adoption Dynamics of Internet Technologies: Models and Case Studies

Abstract
Today, more than any time in history, our life-styles depend on networked systems, ranging from power grids to the Internet and social networks. From shopping online to attending a conference via P2P technologies, the Internet is changing the way we perform certain tasks, which incentivizes more users to join the network. This user population growth as well as higher demand for a better access to the Internet call for its expansion and development, and therefore, fuel the emergence of new Internet technologies. However, many such technologies fail to get adopted by their target user population due to various technical or socio-economical problems. Understanding these (adoption) problems and the factors that play a significant role in them, not only gives researchers a better insight into the dynamics of Internet technology adoption, but also provides them with enhanced guidelines for designing new Internet technologies. The primary motivation of this thesis is, therefore, to provide researchers and network technology developers with an insight into what factors are responsible for, or at least correlated with, the success or failure of an Internet technology. We start by delving deeply into (arguably) the salient adoption problem the Internet has faced in its 40+ years of existence, and continues to face for at least a foreseeable future, namely, IPv6 adoption. The study is composed of an extensive measurement component, in addition to models that capture the roles of different Internet stakeholders in the adoption of IPv6. Then, we extend it to a broad set of Internet protocols, and investigate the factors that affect their adoptions. The findings show performance as the primary factor that not only affected the adoption of IPv6, but also plays a role in the adoption of any other network data plane protocol. Moreover, they show how backward compatibility as well as other factors can affect the adoption of various protocols. The study provides a number of models and methodologies that can be extended to other similar problems in

various research areas, such as network technology adoption and design, two-sided

markets, and network economics.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Electrical & Systems Engineering

First Advisor
Roch Guerin

Keywords
Adoption, Internet, IPv6, Model, Protocol

Subject Categories
Computer Sciences | Electrical and Electronics

ON THE ADOPTION DYNAMICS OF INTERNET TECHNOLOGIES: MODELS AND CASE STUDIES

Mehdi Nikkhah

A DISSERTATION

in

Electrical and Systems Engineering

Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Roch Guérin, Harold B. and Adelaide G. Welge Professor of CS

Graduate Group Chairperson

Alejandro Ribeiro, Rosenbluth Associate Professor of ESE

Dissertation Committee:
Roch Guérin, Professor of CS, Washington University in St. Louis
Santosh S. Venkatesh, Professor of ESE, University of Pennsylvania
Constantine Dovrolis, Professor of CS, Georgia Institute of Technology
Jaudelice Cavalcante de Oliveira, Associate Professor of ECE, Drexel University
Boon Thau Loo, Associate Professor of CIS, University of Pennsylvania

To my beloved parents, Hassan and Zahra,

To my wonderful siblings, Rahi, Leila, Zohreh, and Mani,

and

To my best friend and lovely wife, Sara

Acknowledgments

Earning a doctorate degree and preparing a dissertation is a laborious journey, and is certainly impossible to accomplish single-handedly. Therefore, first and foremost, I would like to express my gratitude to Dr. Roch Guérin, my mentor and dissertation advisor, not only for his input and feedback which made the mere existence of this thesis possible, but also for helping me develop a keen eye for details. Under his guidance I developed critical thinking skills while crafting this piece of work, and for this I am indebted to him.

I would also like to thank Dr. Constantine Dovrolis, with whom I had the pleasure of collaborating on the last part of this dissertation. I am truly grateful to him for his sincere support and suggestions, which greatly affected the finalization of this dissertation. I owe a special note of gratitude to Dr. Santosh S. Venkatesh, the chair of my dissertation committee, who provided me with his thoughtful advice throughout my graduate studies. I am also grateful to Dr. Jaudelice Cavalcante de Oliveira and Dr. Boon Thau Loo for agreeing to serve on my dissertation committee, and for providing thoughtful and constructive feedback.

I am sincerely thankful to my best friend and kindhearted wife, Sara, for being a pillar of strength and encouragement in this journey, supporting me along the way regardless of circumstances.

I am also deeply indebted to my parents and siblings, who were always caring and supportive even though they were thousands of miles away. Their high expectations as well as their constant encouragements paved the road for me to achieve my goals.

Last but not least, I want to sincerely thank my friends, Shahin Shahrampour, Mahyar Fazlyab, Aryan Mokhtari, Solmaz Torabi, Shermin Hassanpour, Behnam Hassanpour, Pooya Naseri Nosar, Tannaz Malekian, and Adam Drescher, who over the last six years brought me joy and comic relief from the enormous pressures of work.

ABSTRACT

ON THE ADOPTION DYNAMICS OF INTERNET TECHNOLOGIES:

MODELS AND CASE STUDIES

Mehdi Nikkhah

Roch Guérin

Today, more than any time in history, our life-styles depend on networked systems, ranging from power grids to the Internet and social networks. From shopping online to attending a conference via P2P technologies, the Internet is changing the way we perform certain tasks, which incentivizes more users to join the network. This user population growth as well as higher demand for a better access to the Internet call for its expansion and development, and therefore, fuel the emergence of new Internet technologies. However, many such technologies fail to get adopted by their target user population due to various technical or socio-economical problems. Understanding these (adoption) problems and the factors that play a significant role in them, not only gives researchers a better insight into the dynamics of Internet technology adoption, but also provides them with enhanced guidelines for designing new Internet technologies. The primary motivation of this thesis is, therefore, to provide researchers and network technology developers with an insight into what factors are responsible for, or at least correlated with, the success or failure of an Internet technology. We start by delving deeply into (arguably) the salient adoption

problem the Internet has faced in its 40+ years of existence, and continues to face for at least a foreseeable future, namely, IPv6 adoption. The study is composed of an extensive measurement component, in addition to models that capture the roles of different Internet stakeholders in the adoption of IPv6. Then, we extend it to a broad set of Internet protocols, and investigate the factors that affect their adoptions. The findings show performance as the primary factor that not only affected the adoption of IPv6, but also plays a role in the adoption of any other network data plane protocol. Moreover, they show how backward compatibility as well as other factors can affect the adoption of various protocols. The study provides a number of models and methodologies that can be extended to other similar problems in various research areas, such as network technology adoption and design, two-sided markets, and network economics.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Today, more than any time in history, our life-styles depend on networked systems, ranging from power grids to the Internet and social networks. From shopping online to attending a conference via P2P technologies, the Internet is changing the way we perform certain tasks, which incentivizes more users to join the network. This user population growth as well as higher demand for a better access to the Internet call for its expansion and development, and therefore, fuel the emergence of new Internet technologies. However, many such technologies fail to get adopted by their target user population due to various technical or socio-economical struggles. The struggle in some instances is more critical than others. For instance, adoption of "Google+", a social network sponsored by Google, is dubbed by many as an utter failure, however, it only affects a company and its revenue, as opposed to adoption of IPv6, which is critical for a sustainable expansion of the Internet and affects its livelihood and future. Understanding these (adoption) problems and the factors that play a significant role in them, not only gives researchers a better insight into the dynamics of Internet technology adoption, but also provides them with enhanced guidelines for designing new Internet technologies. This, results in lower design cost and more successful adoption instances, which by themselves incentivize higher investments in network technologies, and close a positive feedback loop for development of networks, and therefore, guarantee the sustainability of our network-dependent life-style. The primary motivation of this thesis is, therefore, to provide researchers and network technology developers with an insight into what

factors are responsible for, or at least correlated with, the success or failure of a network technology.

The focus of this thesis is predominantly on the Internet, because not only its rapid expansion[1] as well as its association with a vast number of innovative technologies, *e.g.,* web, P2P technologies, social networks, make it a unique network on which we are most dependent, but also also it has experienced a large number of technology adoption instances, and therefore, many more adoption struggles compared to any other network. We can point to many such struggles in the Internet such as the adoption of BGPSec, DNSSec, IPSec, SCTP, and man more similar instances, however, there is one technology that is arguably the salient adoption problem the Internet has faced in its 40+ years of existence, and continues to face at least for a foreseeable future, namely, IPv6 adoption. The importance of IPv6 comes from the role it plays in guaranteeing a sustainable expansion of the Internet in the future, *i.e.,* without IPv6, addressing more than about 4 billion devices is impossible without creating further issues including performance, scalability, etc.

Therefore, the first chapter of this thesis focuses on providing a deep understanding of why and how different factors affected the progress of IPv6 adoption. In particular, it investigates and reports on the evolution of IPv6 adoption, and the factors that affected it in the last two decades. The study is composed of two different components, namely, measurement and modeling. The extensive measurements

[1] See `http://www.internetlivestats.com/internet-users/` for an illustration of the growth of the Internet user population.

done by us, and others, show the progress of IPv6 adoption occurred in distinct phases. They also show the changes in factors affecting IPv6 adoption across various Internet stakeholders. Finally, through a modeling effort, it demonstrates the causal relationship between the changes in those factors, and the phases of progress in IPv6 adoption.

Then, the second chapter investigates a number of future scenarios that can still derail or speed up the progress of IPv6 adoption. Using game theoretic and optimization models, the second chapter provides a framework for understanding the complex interactions between different stakeholders, and the interplay of various factors in their decision making processes. The models identify two different scenarios, namely, one where the decisions of some Internet stakeholders can create an unpredictable ecosystem that can lead to derailing the current progress of IPv6 adoption, and another where a minimal coordination between those stakeholders create a predictable environment. Then, using the models in the second scenario, the factors responsible for speeding up the progress of IPv6 adoption are also identified. While the models are specifically tuned for the IPv6 adoption problem, it can be used by slight changes for future similar adoption instances.

After understanding the adoption dynamics of IPv6, the factors associated with them, and the interaction of various stakeholders, in order to expand the breadth of our study and its applicability, we turn our focus to a broader set of Internet technologies, and in particular, protocols. Therefore, the third chapter of this the-

sis focuses on a broader set of Internet protocols, and investigates, using rigorous statistical methods, their adoption dynamics. In particular, this chapter explores the correlation between success or failure of protocols and a number of characteristics associated with them. The statistical models, *e.g.,* logistic regression, not only confirm the existence of such correlations, but also quantify the effects of each one of those factors that play a role in the success or failure of a protocol. This process is then combined with our engineering intuition, which improves the accuracy of the findings, and provides more intuitive and interpretable outcomes. The methodology developed in this chapter is easily extendable to many other network technology adoption problems, since it is based on collecting relevant data and using well-established statistical tools. Each of these chapters are introduced in greater details in Sections 1.1 to 1.3. The modeling frameworks and methodologies developed in this dissertation are applicable in many different network settings, and can motivate further research in the general area of technology adoption.

1.1 Migrating the Internet to IPv6

The Internet has grown far beyond what its original designers anticipated. As a result and even if the original 32-bit IPv4 addresses may have initially seemed an inexhaustible resource, we have run out of them[2]. The need for a solution was

[2]IANA allocated its last large block of IPv4 addresses in February 2011, and the RIRs are rapidly following suit, *i.e.,* see http://www.potaroo.net/tools/ipv4 for an up-to-date status.

recognized early on and led to the standardization of IPv6 in 1995 [24]. IPv6 boasts a 128-bit address field, and therefore this time a truly inexhaustible address space. However, even if IPv6 was standardized close to 20 years ago and the IPv4 address exhaustion is now a reality, the Internet's migration to IPv6 has been anything but smooth, to the point that many have at times expressed doubts it would ever happen.

Migrating the Internet to IPv6 involves two dependent factors, the availability (and stability) of IPv6 solutions across the Internet infrastructure (from applications to network components), and the adoption (and use) of those solutions by Internet stakeholders. In that context, the goals of this study are two-fold. It seeks, using empirical data gathered over time by us and others, to document and elucidate the progress of the availability and use of IPv6 across major Internet stakeholders (more on this below). It also aims to build and validate a simple model that captures some of the cause and effect relationships that produced major changes in those empirical observations.

Empirical data suggest an evolution that went through roughly three major phases since IPv6 was first introduced. The first phase, from IPv6 inception (circa 1995) until about 2009, is best characterized as stagnant, *i.e.,* IPv6 usage experienced little or no growth even if IPv6 as a technology matured considerably during that time. As we argue later in the chapter, the lack of maturity (compared to IPv4) of initial versions of IPv6 solutions likely contributed to IPv6 limited early

appeal. A second phase followed from 2009 until early 2012, where while IPv6 usage remained mostly marginal, there were telltale signs of its emergence. A third phase started in late 2012, with IPv6 usage slowly accelerating, so that an eventual migration of the Internet to IPv6 now appears likely, albeit still distant.

The study's contributions are in documenting and to some extent revealing the stages IPv6 development and deployment went through across stakeholders. In this study, we also propose a simple model to explicate the cause and effect relationships that have and are driving the Internet's migration to IPv6, and offer qualitative evidence of the model's predictive ability.

1.2 Maintaining the Progress of IPv6 Adoption

IPv6 was designed to address the issue of IPv4 address scarcity, and even though the study in Chapter 2 shows its adoption is accelerating, there are hurdles that can impede or slow down its progress in the future. Although these hurdles are not (anymore) of a technical nature, years of technology disparity between IPv4 and IPv6 caused a marginal adoption of IPv6 across major Internet stakeholders [62], which in addition to incompatibility of the two technologies forced the use of translation mechanisms to allow IPv6-only users access to the IPv4-only Internet. These translation mechanisms are widely used today by ISPs such as CERNET2 in China, and Verizon Wireless and T-Mobile in the U.S. CERNET2 [90] (an academic network), already had over 400k *IPv6-only* users in 2009, is expected to reach 3 million

by the end of 2015 (see [12, 13]), and uses "IVI", which translates IPv4 traffic to IPv6 and vice versa. Similarly, Verizon Wireless and T-Mobile are now primarily relying on IPv6 addresses for new cell-phone subscribers [77,82], and use "NAT444" and "464XLAT" as their translation mechanisms, respectively. While necessary for a transition, the quality degradation those mechanisms introduce [3, 14, 27, 53] reduces motivation for the new users to adopt IPv6. This is an instance of hurdles in front of the progression of IPv6 adoption in the future. Our initial intuition was that besides the above instance, the distributed structure of the Internet can also affect the progression of IPv6 adoption. Specifically, the benefit of migrating to IPv6 depends to a large extent on what others in the Internet do. This is not an uncommon situation (*e.g.,* see [2] for a related discussion in the context of Internet security protocols), but uncertainty in the decisions of others can significantly delay the adoption of a new technology.

A goal of this study is, therefore, to explore and explain strategies that can derail or speed up the current progress of IPv6 adoption. These strategies require careful assessments as we are dealing with a highly decentralized system (the Internet). To better understand the extent to which these strategies can affect IPv6 adoption, several simple yet representative scenarios and models were developed. The focus of these models is on the decision making process of independent and decentralized stakeholders across the Internet, and how those decisions can affect IPv6 adoption. We acknowledge up-front the many simplifying assumptions these models rely on (a

necessity in most modeling efforts), and their lack of completeness. However, they incorporate major aspects of IPv6 adoption decisions, namely, (i) heterogeneity in the Internet stakeholders making decisions; (ii) a representative sample of available technology options; and (iii) the dependencies that exist across decisions.

Our findings from these models indicate that independent decision making process of ISPs can negatively affect IPv6 adoption. In other words, disagreement between ISPs on connectivity option offerings, adds uncertainty to the factors that affect IPv6 adoption decisions of the Internet stakeholders, and makes it hard to identify winning strategies. As a result of this uncertainty, migration to IPv6 slows down, or at the very least becomes haphazard. Another finding of the models is that even minimal coordination among ISPs in offering connectivity options, *e.g.,* an Internet-wide consensus on offering IPv6 as one of the connectivity options, can significantly improve our abilities to identify strategies that hasten the IPv6 migration process. Although consensus alone is not sufficient, it makes it easier for the Internet stakeholders to identify winning strategies that can, at the same time, speed up the migration of the Internet to IPv6.

This study's contributions are, therefore, two-fold:

(i) It shows how distributed decision making of the Internet stakeholders, in the presence of competing solutions to the problem of IPv4 address scarcity, can negatively affect identifying winning strategies, and therefore, contribute to the lingering of the current quandary in IPv6 adoption; and

(ii) It illustrates how the introduction of limited coordination among ISPs, which is not in itself enough for IPv6 success, can help determine the impact of different parameters on IPv6 adoption, and hence, facilitate a smoother migration process.

1.3 Key Factors in Protocol Adoption

Over the past decades, the networking community has learned much about protocol design, be it in terms of performance, scalability, security, etc., or even in some cases guaranteeing the correctness of a protocol. However, we know much less about what controls a protocol's success in the "real world". IPv6 is a well-known instance, which more than two decades after its introduction still struggles to achieve wide adoption. And there are many other examples. Since 1969 the Internet Engineering Task Force (IETF) has produced over 3100 *standards track* Request for Comments (RFCs). However, in spite of a rigorous vetting process, the odds are little better than even[3] for those protocols to succeed, *i.e.,* be widely adopted by their target audience.

This raises a number of important questions that, surprisingly, have not been really addressed by previous research[4]. In particular, are specific features or properties more important than others when it comes to influencing a protocol's success? Clearly, technical correctness is important, but we have arguably made much

[3]A random sample of close to 200 standard tracks RFCs yields a success rate of about 60%.

[4]Related works are primarily in the realm of "network economics" and therefore with a different focus than this study.

progress in weeding out flawed protocols. External factors such as luck or commercial interests will always be present, but are unlikely to translate into systematic biases. The question is whether it is possible to carry out a quantitative and statistically rigorous investigation of protocols and protocol extensions[5] to identify factors with a significant influence on their success (or failure). Additionally, do these factors vary as a function of a protocol's type, *i.e.,* the functionality and environment it targets?

In this study, we apply statistical tools to mine a rich and diverse repository of protocols, namely *standards track* RFCs. Standards track RFCs correspond to protocols that have progressed through rounds of discussions in an IETF Working Group (WG), and been deemed stable and significant enough to warrant formal publication. This should, therefore, eliminate technically flawed protocols, as well as those with little community support. Our goal is to identify statistically significant features that play an important role in a protocol's success, with success defined as "broad-based" adoption among intended users. Note that in identifying such features, we do not seek to build a prediction tool. Instead, we aim to offer guidance to protocol designers by highlighting features that may be of particular significance for different types of protocols.

Our approach is three-prong. We first identify features, which reflect protocol characteristics that *may* play a role in their success. Crafting such a list is a some-

[5]For conciseness and unless otherwise warranted, we use the term protocol to refer to both *new* protocols and extensions or new versions of *existing* protocols.

what subjective process that borrows on our experience with protocols and protocol design. Next, we construct a data set that we analyze statistically. This data set is built from a random sample of standards track RFCs, which are then characterized in terms of their features as well as labeled as successful or not. Finally, we apply a well-established classification framework to extract protocol features that show statistically significant correlation with the success or failure of protocols. The results are then analyzed to explore their implications, and perform limited validation.The outcomes of the analysis are both intuitive and surprising. As expected, prediction accuracy remains in the $70 - 80\%$ ranges, as our focus on design features does not account for the likelihood that other non-technical factors play a role in a protocol's success. The results also confirm that markedly different features affect the success of protocols of different types. For example, while backward compatibility plays a critical role in the success of protocol extensions or new versions of existing protocols, it is of little relevance when it comes to new protocols. Other findings are, however, less immediately obvious. For example, the most significant factor contributing to the success of new application and transport layer protocols was the extent to which they were of benefit to other existing protocols. Similarly, the success of network control protocols depended heavily on their ability to realize their full value once deployed within a domain (as opposed to Internet-wide deployment).

1.4 Research Publications

The research work presented in this dissertation has been published in several leading conferences and journals An initial version of our work on migration of IPv6, titled "Assessing IPv6 Through Web Access - A Measurement Study and Its Findings" (co-authored with Roch Guérin, Yiu Lee, and Richard Woundy) [63] was published in the proceedings of ACM CoNEXT, in Tokyo, Japan, on December 6-9, 2011. A more mature version of the work including measurements from an extended period of time and a simple model, is accepted for publication as "Migrating the Internet to IPv6: An Exploration of the When and Why" (co-authored with Roch Guérin) [59] to appear in IEEE/ACM Transactions on Networking.

The work on modeling the future scenarios of IPv6 adoption was first published as " Migrating to IPv6 - The Role of Basic Coordination" (co-authored with Roch Guérin) [61] in the proceedings of IFIP Networking, in Trondheim, Norway, on June 2-4, 2014. The work was extended by including robustness tests, and other scenarios, and is currently under review in one of the leading networking journals.

Our results on the last topic titled "Why didn't my (great!) protocol get adopted?" (co-authored with Constantine Dovrolis and Roch Guérin) is currently under review in a top-tier workshop in networking, and plans for a journal version of the work is under consideration.

Chapter 2

Migrating the Internet to IPv6: An Exploration of the When and Why

2.1 Introduction

The Internet has grown far beyond what its original designers anticipated. As a result and even if the original 32-bit IPv4 addresses may have initially seemed an inexhaustible resource, we have run out of them[6]. The need for a solution was recognized early on and led to the standardization of IPv6 in 1995 [24]. IPv6 boasts a 128-bit address field, and therefore this time a truly inexhaustible address space. However, even if IPv6 was standardized close to 20 years ago and the IPv4 address exhaustion is now a reality, the Internet's migration to IPv6 has been anything but smooth, to the point that many have at times expressed doubts it would ever happen.

Migrating the Internet to IPv6 involves two dependent factors, the availability (and stability) of IPv6 solutions across the Internet infrastructure (from applications to network components), and the adoption (and use) of those solutions by Internet stakeholders. In that context, the goals of this chapter are two-fold. It seeks, using empirical data gathered over time by us and others, to document and elucidate the progress of the availability and use of IPv6 across major Internet stakeholders (more on this below). It also aims to build and validate a simple model that captures some of the cause and effect relationships that produced major changes in those empirical observations.

[6]IANA allocated its last large block of IPv4 addresses in February 2011, and the RIRs are rapidly following suit, *i.e.,* see http://www.potaroo.net/tools/ipv4 for an up-to-date status.

Empirical data suggest an evolution that went through roughly three major phases since IPv6 was first introduced. The first phase, from IPv6 inception (circa 1995) until about 2009, is best characterized as stagnant, *i.e.,* IPv6 usage experienced little or no growth even if IPv6 as a technology matured considerably during that time. As we argue later in the chapter, the lack of maturity (compared to IPv4) of initial versions of IPv6 solutions likely contributed to IPv6 limited early appeal. A second phase followed from 2009 until early 2012, where while IPv6 usage remained mostly marginal, there were telltale signs of its emergence. A third phase started in late 2012, with IPv6 usage slowly accelerating, so that an eventual migration of the Internet to IPv6 now appears likely, albeit still distant.

The chapter's contributions are in documenting and to some extent revealing the stages IPv6 development and deployment went through across stakeholders. The work in this chapter also proposes a simple model to explicate the cause and effect relationships that have and are driving the Internet's migration to IPv6, and offers qualitative evidence of the model's predictive ability.

The rest of the chapter is organized as follows. Section 3.7 briefly reviews relevant prior works. Section 2.3 introduces Internet stake-holders and their respective roles, and reports their use of IPv6 over time. Section 2.4 identifies factors that likely affect the decisions of Internet stakeholders when it comes to IPv6, and discusses the impact that changes in these factors may have had. Section 2.5 introduces a simple model to capture these decision processes, and uses it to qualitatively re-

produce the trends reported in the data of Section 2.3. Section 2.6 summarizes the work's contributions.

2.2 Related Works

The Internet's transition to IPv6 has been extensively studied, and we only review a sample of representative works, some of which are detailed further in the next section. Most works fall in either one of two major categories: measurement (empirical) or modeling (analytical) studies.

Empirical studies have sought to measure IPv6 availability and performance at both an Internet-wide scale and by focusing on individual components. See for example [64] for a useful albeit slightly dated overview of the status of IPv6 across the Internet, or CAIDA [11, 26] that arguably offers one of the more comprehensive repository of related information. Other studies have focused on quantifying adoption across Autonomous Systems (ASes) [31, 68], among end-users [32], and in Operating Systems (OSes) [18, 38]. Performance issues in OSes have been explored in [58, 91], while investigations aimed at end-to-end performance have compared IPv4 and IPv6 using metrics such as path delay and packet loss [86, 93]. On the modeling front, many studies have sought to formulate the IPv6 adoption question in the context of an economic framework, in an attempt to capture the many interacting factors affecting it [30, 34, 62, 79].

Finally, a recent comprehensive investigation of the status and progress of IPv6

prevalence across the Internet ecosystem was reported in [22]. It measured changes in address allocation, DNS readiness, routing, etc., and is closest in motivations to this chapter. An important difference though is in our attempt to develop a model that can explain some of the measurement results on which we report. In particular, this chapter combines measurements and models to not only document, but also to some extent explain the evolution of IPv6 adoption.

2.3 Quantifying the Internet's Migration to IPv6

This section reports on the evolution of IPv6 "adoption" across Internet stakeholders. Those stakeholders are diverse and adopting IPv6 has very different meanings across them. Hence, it is useful to first describe them, together with what IPv6 adoption means for each. This is the goal of the next sub-section, which also introduces how IPv6 adoption is measured.

2.3.1 Internet Stakeholders

There has been much interest for what drives the Internet's growth and the roles its stakeholders play, *e.g.,* as demonstrated by the creation of an OECD Working Group on Internet Governance[7]. A recent report [47] offers an initial taxonomy of Internet stakeholders that lists, among others, Internet Technology Developers (ITDs), Internet Service Providers (ISPs), Internet Content Providers (ICPs), and

[7] www.oecd.org/internet/broadband/oecdresourcesoninternetgovernance.htm.

Internet (content) consumers or users. We focus on those, as they are the major actors in the Internet's migration to IPv6, and review their roles and how to best quantify their migration to IPv6. This is followed by the presentation of measurement data, gathered by others and us, that offers a timeline for the Internet's migration to IPv6.

ITDs

They build the technologies behind the Internet, and are, therefore, necessary precursors to any new Internet capability, including IPv6. In other words, they develop and release IPv6 versions of their products that are then deployed by other stakeholders to realize an IPv6 Internet. Hence, measuring IPv6 "adoption" among ITDs calls for tracking the availability *and* stability of IPv6-capable products (an IPv6 version may be available, but until it is as stable as its IPv4 counterpart, it is unlikely to be widely adopted).

ISPs

They provide (Internet) connectivity to users and ICPs through equipment purchased from ITDs. Their adoption of IPv6 is through upgrading their infrastructure to IPv6, *i.e.,* by supporting routing and forwarding of IPv6 traffic. This adoption can be measured in a number of different ways, but we rely on two representative metrics: (i) the number of major transit Autonomous Systems (ASes) that advertise IPv6 capabilities; (ii) the number of (peering) links that exist between them.

The first offers insight into the overall penetration of IPv6 among ISPs, while the latter captures the density of IPv6 connectivity (both affect end-to-end connectivity quality).

ICPs

They own the content that makes up for much of the Internet's value (to users). For an ICP, IPv6 adoption implies native IPv6 access to its content. This requires upgrades to its local infrastructure (or that of its hosting provider), and advertising IPv6 accessibility (through DNS) to users. Measuring IPv6 accessibility among ICPs, therefore, calls for tracking which ICPs advertise IPv6 addresses. ICPs, like ISPs, are, however, diverse in size and popularity, and accounting for those differences can offer a more accurate perspective on IPv6 adoption. In the next section, we report measurements of both overall IPv6 adoption by ICPs, as well as based on their popularity. We use the latter to later estimate the volume of IPv6 traffic contributed by ICPs.

Users

Users derive "value" from accessing content, use of Internet services, etc. They are mostly oblivious to technology choices, but their expectations for the underlying technology have implications for IPv6: **(i)** IPv6 applications should be available and stable; **(ii)** IPv6 connectivity should be on par with IPv4; and **(iii)** content should be accessible over IPv6. Hence, using IPv6 addresses for new users once IPv4

addresses have been exhausted, is feasible only if those conditions are met. Because a comprehensive census of IPv6 users is not feasible, we measure IPv6 "adoption" among users using statistical estimates based on representative samples.

2.3.2 Assessing the Internet's Migration to IPv6

This section presents empirical data on the evolution of IPv6 adoption among ITDs, ISPs, ICPs, and users. As mentioned before, the data points to a three-phase adoption:

Phase I $[1995 - 2009]$: Stagnation;

Phase II $[2009 - 2011]$: Emergence;

Phase III $[2011-)$: Acceleration.

We provide next evidence in support of this conclusion.

ITDs

There are many technologies involved in delivering Internet connectivity. We focus on IPv6 progress for a representative subset, namely, routers/switches, Operating Systems (OSes), and applications.

Router/Switch Manufacturers Support for IPv6 came in early in routers/switches, *e.g.,* between 1998 to 2000, when Juniper introduced its first series of IPv6-capable routers [46]. Cisco quickly followed suit by introducing IPv6 capability in CISCO IOS routers and L3 switches. Early availability, however, did not equate qual-

ity/stability. In particular, a 2007 study [93] showed that IPv6 forwarding plane lagging behind its IPv4 counterpart, with routers/switches the primary culprits. Those early stability problems are, however, now over and a study we conducted in 2011 [63] showed that the IPv6 and IPv4 forwarding planes now perform similarly.

OS Developers Support for IPv6 appeared first in Linux 2.1.8 (in 1996), but remained in experimental status until around 2005. Microsoft Windows 2000 did support IPv6, but not by default, and Microsoft did not ship Windows OSes with default IPv6 support until 2007 (Windows Vista). Apple introduced IPv6 by default in 2003 in Mac OS X v10.3. As with routers, early IPv6 offerings were plagued by problems, *e.g.,* [91]. Performance, however, improved over time across all operating systems, *e.g.,* [58] showed in 2009 that IPv6 and IPv4 performance were on par in Microsoft Windows Vista and in Linux Ubuntu. As of today, IPv6 is available in nearly all operating systems [43] (Windows Phone 8 added support for IPv6 in 2011, and Android with its Lollipop release) with few if any remaining performance problems [38].

Internet Application A comprehensive list of IPv6 capable applications (with their IPv6 launch date/version) is available at [41,42]. In this subsection, we rely on a set of popular applications, to gauge the evolution of applications' IPv6 readiness. Table 2.1 gives the launch date of the IPv6 version of applications in this target set (or NA when an IPv6 version is not yet available). The table indicates a slow but

steady progress in adding IPv6 support from the late 1990's until today. However, as with routers, switches and OSes, IPv6 support was not always synonymous with stability or quality. Consequently, many applications initially shipped with IPv6 disabled by default, and some still do (*e.g.,* VMware vSphere ESX/ESXi had IPv6 disabled by default until v. 5.1 [85]).

In summary, after a relatively slow start, IPv6 support is now readily available across all major Internet technologies. Maturity and stability of those offerings is, however, relatively recent. The lack of stability in early versions may partially explain some of the findings on which we report next, namely, a relative stagnation of IPv6 adoption among other Internet stake-holders (ISPs, ICPs, and users) in IPv6 early years.

IPv6 status across ISPs

As ISPs upgrade their network to IPv6 and advertise it to other ISPs, they affect overall IPv6 Internet connectivity. To measure this impact, we focus on IPv6 adoption among "major transit ISPs" that carry a large share of Internet traffic[8]. CAIDA has been conducting such a study since 2005 [11], tracking all major IPv4 and IPv6 transit ASes and their peering links. We summarize in Table 2.2 some of CAIDA's more salient results, which illustrates the evolution of IPv6 adoption among major transit ASes.

[8]Recent reports have identified shifts in Internet traffic patterns [51], but those ISPs continue to be responsible for the bulk of the Internet traffic.

Applications	**IPv6 Launch**
Internet Explorer	1999
Microsoft DNS	2000
SSH	2000
Java	2002
Apache HTTP Server	2002
Python	2003
Microsoft SQL Server	2005
Open VPN	2006
Windows Media Player	2006
Mozilla Thunderbird	2007
Microsoft Outlook	2007
Safari	2007
MySQL	2008
Mozilla Firefox	2008
Samba	2008
VMWare vSphere ESX	2011
Office 365	2013
Skype	NA
Xbox 360	NA

Table 2.1: IPv6 launch of key Internet applications (from [41, 42, 56, 85])

	# of ASes (IPv4)	# of ASes (IPv6)	% of ASes (IPv6)	# of Peering Sessns. (IPv4)	# of Peering Sessns. (IPv6)	# of IPv6 Peering per 100 IPv4 Peering
2009	23k	515	2.24%	50k	1904	3.8
2011	29k	1183	4.08%	78k	2738	3.51
2013	34k	2419	7.11%	109k	8881	8.15

Table 2.2: Internet AS core Evolution (Data from CAIDA's website [11]).

CAIDA's data show that by 2009 barely 500 or just over 2% of the major transit ASes were IPv6 capable. The next two years, $2009-2011$, hint at the beginning of a transition with a doubling of this number to 1183, with IPv6 penetration itself also nearly doubling. This trend continued, and the number of IPv6 capable transit ASes again doubled by late 2013. This indicates that in spite of a slow start, a critical threshold seems to have been crossed, with IPv6 deployment now expanding rapidly. The progression of the number of IPv6 peering links/sessions (a measure of IPv6 connectivity) displays a similar trend (last three columns of Table 2.2). Ripe Labs carried out a similar study [68] recording all ASes (from transit to edge, including content ISPs) advertising at least one IPv6 route, which yielded results consistent with the three-phase progression of CAIDA's more focused data.

ICPs IPv6 Accessibility

Internet content is accessible (to users) in many different forms, but websites host the vast majority of it. Tracking IPv6 accessibility across public websites, therefore, offers a reasonable estimate of IPv6 deployment among content providers. Early (circa 2004) estimates [86] reported that barely $1,000$ out of more than 51 millions websites (see `http://goo.gl/Ydql7U`) were IPv6 accessible, *i.e.,* a negligible fraction (less than 0.001%). This confirms the limited appeal of IPv6 in those initial years.

In 2009, we started an independent set of measurements, tracking IPv6 accessibility of Alexa's top one million websites[9] (see `http://www.alexa.com`). Alexa's top one million websites span a broad range of categories (commercial, educational, entertainment, etc.) and popularity, and offer a representative sample of the now more than 1 billion websites in existence (as of September 2014 based on www.internetlivestats.com). The methodology behind our measurements is documented in [63], but essentially consists of three steps: **(i)** issuing DNS queries for websites in the list; **(ii)** downloading the homepage of websites for which DNS returned both "A" and "AAAA" records, *i.e.,* websites with IPv4 and IPv6 addresses[10]; **(iii)** time-stamping and recording the results in a database.

[9]The list now includes over 8 millions websites, because of churn in Alexa's top 1 million list and additions from local DNS caches.

[10]Websites accessible only over IPv6 represent only a minute fraction of monitored websites, and are, therefore, ignored in the measurements.

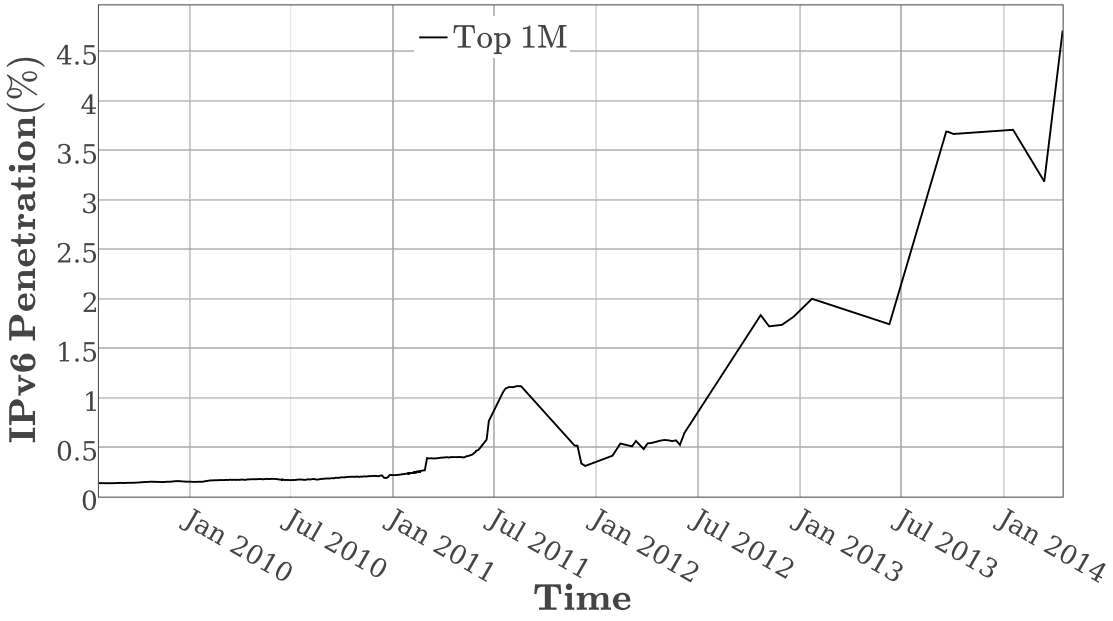


Figure 2.1: IPv6 Adoption among Alexa's Top 1M Websites.

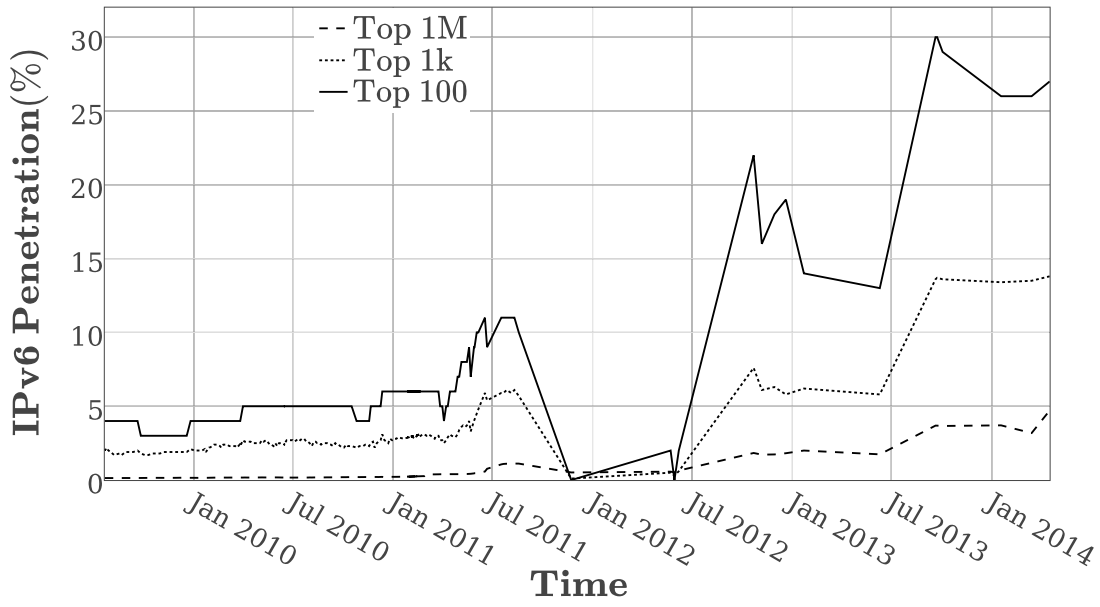


Figure 2.2: IPv6 Adoption among the Top 100, Top 1k, and Top 1M Websites.

Fig. 2.1 reports the results of our measurements, showing that while IPv6 adoption remained low in the $2009 - 2011$ period, it improved on its earlier marginal adoption (it grew from essentially 0 to about 0.2% by early 2011). A momentous change appears to have occurred in 2011, likely spurred by the "official" exhaustion of IANA's IPv4 address pool in February 2011, and by a greater awareness contributed by events such as the World IPv6 Day (`http://www.worldipv6day.org`) that produced a large albeit somewhat transient increase in mid 2011. The temporary gains of the World IPv6 Day were eventually cemented after the World IPv6 Launch in mid 2012 (`http://www.worldipv6launch.org/`), with IPv6 adoption transitioning to a faster pace (approximately doubling every year), and reaching a penetration close to 5% by mid 2014. This roughly mirrors the trend observed for ISPs.

Fig. 2.2 expands the view of Fig. 2.1, showing IPv6 accessibility as a function of a website's *popularity, i.e.,* it reports separately IPv6 accessibility for the top 100, top 1000 and top 1 million websites. The figure clearly illustrates that more popular websites are more likely to be IPv6 accessible (by as much as a factor 6), though all categories follow similar trends.

Estimating the IPv6 User Base

Evaluating the extent to which users have IPv6 connectivity is a challenging problem, not only because of the size of the user population, but also because of the

diversity in how that connectivity is used when available, *e.g.,* many OSes are configured to prefer IPv4 over IPv6 when both are available [80]. Furthermore, changes usually happen at a coarse granularity, *e.g.,* because of an ISP's conversion[11], large scale monitoring is important. For that purpose, we rely on data gathered by Google. Google sites see billions of accesses daily from across the globe, and can monitor how many were over IPv6 [32]. Google's data may under-sample regions such as China where popular alternatives to its services exist, but it nevertheless offers a reasonable assessment of the evolution of the IPv6 user base worldwide.

Google's data show that by 2009, barely 0.2% of users were accessing its services over IPv6. This grew to 0.3% over the next two years, after which growth started to accelerate to reach 3% by early 2014 (a ten-fold increase). This roughly matches the three phase growth pattern of ISPs and ICPs.

The next section seeks to develop a better understanding for the reasons behind this three-phase adoption pattern.

2.4 IPv6 Ecosystem

Explicating the evolution of the Internet's migration to IPv6 calls for a better understanding of what drives Internet stake-holders to adopt IPv6 in the first place. In other words, what factors affect those decisions and how? Users are mostly

[11]See for example, T-Mobile's recent announcement [77] to use only IPv6 for users with Android 4.4 KitKat phones.

oblivious to what technology is used to connect them to the Internet, *i.e.,* IPv4 or IPv6, and their choices are typically dependent on decisions made by ITDs, ISPs and ICPs. As a result, we focus on these latter three stake-holders.

All three are complex decision makers, so that modeling their decisions unavoidably involves simplifications. A common approach is to rely on an objective or utility function that (rational) decision makers then seek to maximize [54]. Utility functions vary across stake-holders, but typically incorporate factors such as cost and quality of a product, its value, how widely it is adopted, etc. In this section, we first posit a number of factors and their influence on the decisions of ITDs, ISPs, and ICPs. We then identify and characterize changes in those factors, and establish how they may have produced the three-phase migration process documented in the previous section.

2.4.1 Decision Factors

In identifying factors and their role in the (IPv6) "adoption" decisions of ITDs, ISPs and ICPs, we consider each separately.

ITDs

They develop IPv6 versions of their products based on expectations of *demand* for those products. As alluded in Section 2.3.2, this demand, however, depends on the availability of those very same products (IPv4 versions were, at least initially,

a perfect substitute); in the process creating a chicken-and-egg problem that may have contributed to their slow maturation. The problem is compounded by the fact that availability alone is not sufficient. Because IPv4 serves as a benchmark to which IPv6 is compared, the quality and stability of IPv6 products affects demand; in particular by ICPs whose revenues are affected by the quality of users' experience. Formalizing the impact of those dependencies in the context of a simple model is the subject of Section 2.5.

ISPs

The growing scarcity of IPv4 addresses is the primary motivation for an ISP to adopt IPv6. This, however, calls for upgrading its network and operational practices. This one time cost can result in an ISP deferring such a decision, especially since alternatives exist for dealing with IPv4 address shortages. Those include private IPv4 addresses, or securing additional public IPv4 addresses, *e.g.,* through "markets" that are emerging to meet such a demand (see Section 2.4.2).

Large-scale use of private IPv4 addresses has many drawbacks, including the need to deploy "Carrier Grade NATs" (CGNs) or NAT444, and more importantly offers little long-term strategic value (see [45] for a related discussion). Purchasing (new) public IPv4 addresses avoids those problems, but has a cost of its own. One that will likely increase as the supply of available public IPv4 addresses dwindles. This is in part why ISPs such as CERNET2 in China, and Verizon Wireless and

T-Mobile in the U.S., opted to start using IPv6. CERNET2 [90] (an academic network), already had over 400k *IPv6-only* users in 2009, and is expected to reach 3 millions by the end of 2015 (see [12,13]). Similarly, Verizon Wireless and T-Mobile are now primarily relying on IPv6 addresses for new cell-phone subscribers [77, 82].

An ISP's decision to adopt IPv6 and start assigning IPv6 addresses to its users will, therefore, largely depend on the tension between upgrade costs and the cost of procuring new public IPv4 addresses once it exhausts its current pool. The simple model of Section 2.5 seeks to capture this tension.

It should be noted though that adopting IPv6 has implications beyond allocating IPv6 addresses to new users. In particular, users not assigned a public IPv4 address need some form of "translation" service to connect to the public IPv4 Internet. For example, CERNET2 and T-Mobile use IPv6-to-IPv4 translation mechanisms called IVI and 464XLAT, respectively. Verizon Wireless, on the other hand, assigns both private IPv4 and IPv6 addresses to users. A user's IPv6 address is used to connect to IPv6 accessible destinations, while connectivity to the public IPv4 Internet is through the user's private address and NAT444 devices. Translation requirements will, however, eventually disappear once the Internet is fully IPv6 accessible. Hence, while ISPs will incur translation costs after exhausting their public IPv4 addresses, these costs alone are unlikely to play a major role in their decision to upgrade to IPv6.

ICPs

They are mostly oblivious to how their content is accessed, *i.e.,* whether over IPv4 or IPv6, and mostly concerned with how access may affect their revenue. ICPs derive revenues from users in a variety of ways [67], from a user's number of clicks (*e.g.,* Google), to a user's purchasing a good (*e.g.,* Amazon), to how much time a user spends consuming content (*e.g.,* Facebook), etc. In spite of their diversity, these have in common that they are impacted by connectivity *quality* (see [9] for an investigation with a "per-click" revenue model, and [75] for a general study of how a site's "speed" affects conversion rates). Performance of IPv6 users is well known to be negatively affected by translation [3, 14, 27, 53], which can then provide an ICP with the motivation to become IPv6 accessible. This is, however, predicated on IPv6 connectivity being of sufficient quality, and on the number of IPv6-only users being high enough to justify the change and its cost. The former is well illustrated by the "white-listing" [33] that content providers such as Google rely on to control IPv6 connectivity to their content (IPv6 connectivity is allowed only if its quality is on par with that of IPv4). The latter depends on both the expected growth in the number of IPv6 users and on the cost of upgrading the ICP's infrastructure and operational procedure (or those of its hosting provider) to IPv6. This cost is usually proportional to the size of the ICP.

There are clearly other factors that can contribute to an ICP's decision to become IPv6 accessible, *e.g.,* greater ease of obtaining Provider Independent (PI) IPv6

Factors ↗	*Impact on Utility*		
	ISPs	**ICPs**	**ITDs**
Demand for IPv6 Tech.	✗	✗	⊕
IPv4 Address Cost	⊖	✗	✗
Upgrade Costs	⊖	⊖	✗
Translation Cost	∼	✗	✗
# of IPv6 Users	∼	⊕	∼
# of IPv6 ICPs	⊕	∼	∼
IPv6 Quality	✗	⊕	✗

Table 2.3: Effect of increases in IPv6 adoption factors.

addresses [49]. However, improving connectivity quality, and consequently revenue, is one factor common to all ICPs. In contrast, the ability to, say, obtain a PI IPv6 address is attractive only to ICPs without a PI IPv4 address, and this is a relatively small fraction (a random sample of 100 websites in the top 1 million showed that about 80% of them already had a PI IPv4 address[12]). Hence, we expect IPv6 connectivity quality and its impact on ICPs revenue to be a major factor in their decision to become IPv6 accessible.

[12]An exhaustive census is challenging, as accurately verifying address ownership is complex and involves manually cross-checking multiple databases.

2.4.2 Ecosystem changes

Section 2.3 documented changes over time in IPv6 adoption by Internet stakeholders, while Section 4.2 put forward factors that are likely to shape their decisions. In this section, we investigate the extent to which those factors evolved over time, and whether those changes can explain the three phases migration pattern we observed.

As a prelude to this investigation and as a means to classify the impact of the different factors identified in Section 4.2, we record them in Table 2.3 according to how increases ($\nearrow$) in each one of them affect decision makers. The ($\oplus$) and ($\ominus$) symbols in the "Effect" columns indicate whether an increase has a positive or negative impact on a stakeholder's utility. Conversely, an ✗ symbol signals that the factor does not affect the stakeholder's utility, while a $\sim$ indicates that the factor should only have a marginal impact.

Demand for IPv6 Technologies

It is not easy to quantify the demand for IPv6 technologies. However, anecdotal evidence points to near-zero demand in 1995 (the birth of IPv6), followed by government mandates providing some initial impetus in the late 90's, before the looming scarcity of IPv4 addresses became more apparent and resulted in a substantial demand today, *e.g.,* in 2014 Verizon Wireless proceeded to allocate IPv6 addresses to over 45% of its approximately 90 millions subscribers [83]. ITDs likely responded to or anticipated those changes, which may explain the progressive mat-

uration of IPv6 core technologies in the 1990's, followed by the rapid expansion of IPv6 enabled end-devices and applications in the late 2000's.

IPv4 Address Cost

As mentioned earlier, although IANA and most RIRs have run out of IPv4 address blocks to allocate, this does not mean that all public IPv4 addresses are in use. As a matter of fact, recent studies [23, 35] estimate that of the order of about 30% of all public IPv4 addresses are still available (unused). As a result, mechanisms, *e.g.,* markets, have started to appear to facilitate access to those unused addresses. Specifically, following the purchase in 2011 of Nortel's IPv4 addresses by Microsoft at a cost of about $11 per address (`http://goo.gl/ZIA18`), several private markets have emerged such as the IPv4 Market Group (`http://ipv4marketgroup.com`) and IPv4Auctions.com. Both report a steady stream of IPv4 addresses sales at prices ranging from $7 to $18 in 2013 and 2014, with larger blocks, *i.e.,* /15's and /16's having typically lower per address costs than smaller blocks[13].

The role of those markets in facilitating the exchange of IPv4 addresses notwithstanding, their biggest impact is likely to signal to ISPs that IPv4 addresses are not *free* anymore. As Table 2.3 highlights, having to potentially pay for what used to be a free resource, negatively affects an ISP's utility. Expectations that acquiring new IPv4 addresses will become increasingly expensive contribute to strengthening the benefits of adopting IPv6. In the process enticing more ISPs to embark on such

[13]See `http://goo.gl/pDI4gQ`, `http://goo.gl/udHdW1`, and `http://goo.gl/4RCEw9`.

an upgrade, and increasing demand for IPv6 products.

Infrastructure Upgrade Costs

They affect both ISPs and ICPs, and we review each in turn.

ISPs: Upgrading an ISP's infrastructure to support IPv6 is no small task. It involves equipment and operational upgrades, and as can be expected [44], has a cost proportional to the size and complexity of the ISP's infrastructure. As recommended in [44], this cost can be spread out and incurred as part of routine upgrade cycles. Any such upgrade will, however, be more challenging/costly if IPv6 versions of new products are not stable. This introduces a direct coupling between the demand for IPv6 products and the level of investment (by ITDs) required to ensure a sufficient quality. In particular, low investments in IPv6 products because of low anticipated demand result in lower quality products, which in turn drives demand down. Strategic investment decisions by ITDs (or spurred by government mandates) can break the cycle, and trigger an initial demand that will in turn fuel further investment and growth in product quality and eventually demand. The model we introduce in Section 2.5 incorporates this coupling.

ICPs: Upgrading an ICP's infrastructure to IPv6 shares many of the same features as upgrading an ISP's infrastructure to IPv6. As with ISPs, labor and hardware/software equipment costs are the main contributors [44]. Those costs will typically increase with the size of the ICP's infrastructure, and decrease as IPv6

technology matures. Hence, ICPs face a trade-off between delaying upgrading until the technology is sufficiently stable, but then having to perform such an upgrade for a larger user-base. Section 2.5 introduces a simple model that reflects this trade-off and mimics an ICP's decision process.

Translation Costs

As discussed earlier, translation is required for both IPv6 and private IPv4 addresses to allow connectivity to the public IPv4 Internet, and in particular ICPs. Translation costs are roughly proportional to the volume of traffic that needs to be translated [37]. For IPv6, this depends on both the number of IPv6 users and the number and popularity of ICPs requiring translation (not IPv6 accessible). We performed a crude assessment of the evolution of translation traffic based on estimates for the growth in the number of IPv6 users of Verizon Wireless[14] (available at `http://labs.apnic.net/ipv6-measurement/AS/2/2/3/9/4/`) and the number and popularity of ICPs that are not IPv6 accessible. The results are in Fig. 2.3 and assume that all users request the same amount of traffic and that traffic originates from ICPs in proportion to their popularity according to an exponential distribution. The latter is based on measurements for the top $10,000$ websites of a large cable provider. The figure shows that even if recent growth in the number of IPv6-only users contributed to an increase in translation traffic, this volume remains small (less than 0.4% of Verizon Wireless traffic).

[14]Those users do not have a public IPv4 address.

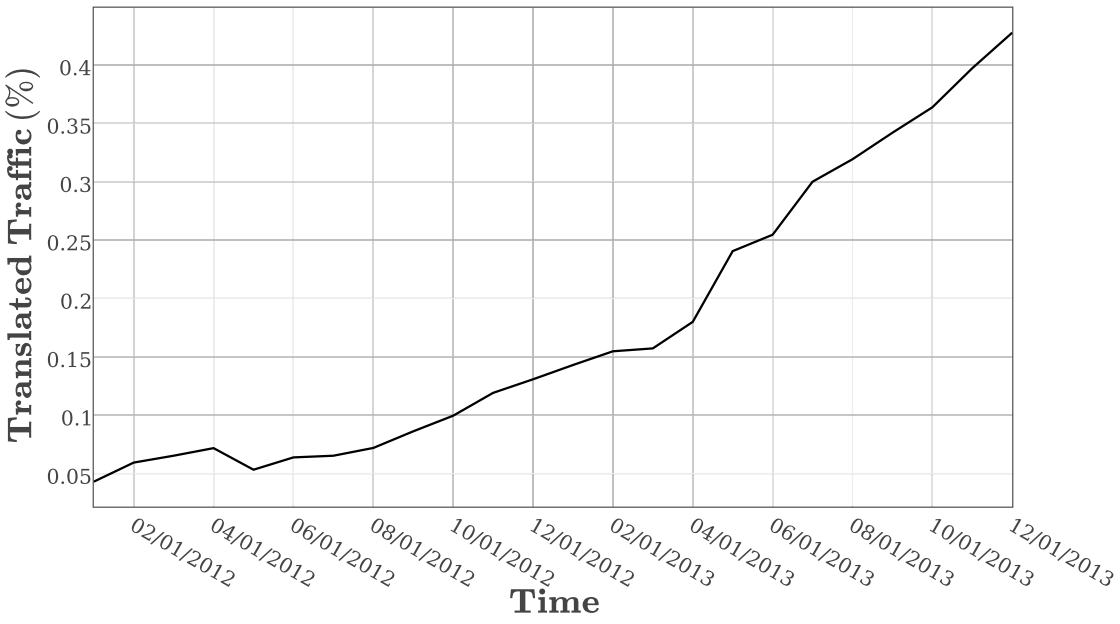


Figure 2.3: Approximating translated traffic volume over time.

Number of IPv6 Users

As seen in the Google data of Section 2.3, the number of IPv6 users has been steadily increasing. This trend is echoed in various public reports pointing to faster IPv6 growth, especially in the Asia-Pacific region[15], where the scarcity of public IPv4 addresses is more severe. A larger IPv6 user base should entice more ICPs to become IPv6 accessible, which would reduce the need for translation and in the process make IPv6 more attractive to ISPs. These positive dependencies could trigger a virtuous adoption spiral of the kind we appear to be witnessing in what we termed Phase III. The model of Section 2.5 offers a possible option for formalizing

[15]See `http://goo.gl/ZG41fU` for a recent announcement.

these dependencies.

Number of IPv6-accessible ICPs

The measurement results of Section 2.3 indicate a strong recent uptick in the number of IPv6 accessible ICPs, which, if the trend persist, should further strengthen IPv6 momentum.

IPv6 connectivity quality

This is the last and possibly most important change in the IPv6 ecosystem, namely, the coming of age of IPv6 when it comes to technology maturity. This maturity manifests itself through improvements in both stability and performance; improvements that finally allowed IPv6 to be on par with IPv4 and in some cases better. We illustrate this in Fig. 2.4 that reports the results of a measurement study started in 2009 (see [63] for details). The study compares IPv6 and IPv4 web download speeds from several vantage points for a large set of websites (including all top 1M sites).

Fig. 2.4 displays the fraction of web servers accessible over both IPv6 and IPv4 for which IPv6 download speed was equal or higher than with IPv4. The figure shows a period of continuous improvement until early 2013, at which point IPv6 was at least as good as IPv4 80% of the time. The remaining gap of 20% is comparable to that of IPv4, which also lags behind IPv6 for 20% of websites. This is not unexpected when comparing two (now) mostly equivalent technologies, where small

configuration or load differences can easily result in one outperforming the other. This hypothesis was confirmed by verifying that when IPv6 is strictly better than IPv4, and vice-versa, the difference in performance is small, *i.e.,* in the range 5 to 10 kbytes/sec.

The results demonstrate that, as of 2013, IPv6 and IPv4 are mostly on par performance-wise. This is undoubtedly the product of improvements made by ITDs. However and interestingly, greater technology maturity is not the only factor behind this change; adoption decisions by ISPs also played a role. In order to better understand this, it is useful to take a closer look at the different components that affect connectivity quality. Specifically, end-to-end connectivity is affected by both end-systems and the network. We proceed next to drill down on each one of these components.

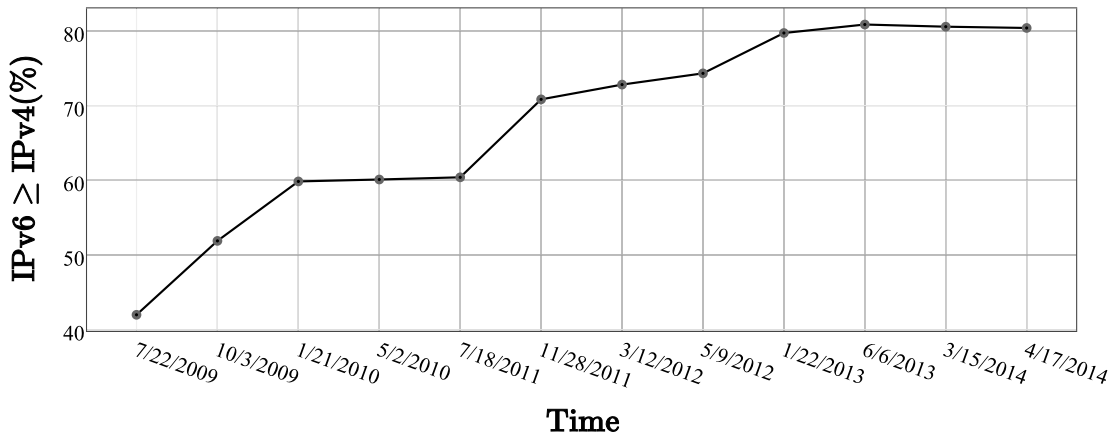


Figure 2.4: Fraction of websites with better or equal IPv6 connectivity than IPv4.

End-Systems IPv6 support in end-systems is dominated by decisions from ITDs[16], *i.e.,* when do they first make it available and how quickly do they ensure that the new software is stable. As reported in Section 2.3.2, IPv6 availability was uneven across OSes with support and improvement across many platforms happening as late as 2009. However, IPv6 support is now stable across all OSes, so that their IPv6 performance is not of concern anymore.

The Network IPv6 network performance depends on both routers' ability to forward IPv6 packets (the data plane), as well as how the path connecting the source to the destination is chosen (the control plane). The first factor depends solely on decisions by ITDs, *i.e.,* their ability to release product upgrades that deliver identical packet forwarding performance in IPv6 and IPv4. The second factor, although clearly affected by ITDs' decisions, is also, as we discuss below, very much dependent on adoption decisions made by ISPs.

There is no denying that IPv6 data plane performance was initially lagging behind that of IPv4. A 2007 study [93] identified a non-trivial gap in end-to-end performance between IPv6 and IPv4, and assigned most of the blame to the data plane. In 2009, we started an independent measurement study aimed at assessing the extent to which this performance gap still existed, and what contributed the most to it. The study involved multiple sources (clients) geographically distributed

[16]Even if the option to turn IPv6 on or off is available, most users will stay with the manufacturer's default configuration.

around the world, which continuously probed over a million websites (including Alexa's top 1M) for IPv6 access, and for those accessible over both IPv4 and IPv6 measured their respective web access performance (download speeds). The study's methodology and its results are documented in [63]. It showed that while as of 2011 a performance gap remained, it was not anymore caused by differences in data plane performance. Instead, control plane factors, *i.e.,* routing and peering decisions affecting IPv6 paths, were the main contributors.

The determination that the IPv6 data plane had finally achieved performance parity, and conversely that control plane factors were now primarily responsible for the remaining performance gap, involved a two step analysis of the available measurement data:

	Top 100k Sites	Top 1M Sites
Same Path Destinations	94%	90%
Diff. Paths Destinations	70%	74%

Table 2.4: IPv6 better or equal to IPv4 between 2009-2011.

Step 1 focused on instances of end-to-end connectivity for which IPv4 and IPv6 made identical control plane decisions, *i.e.,* IPv4 and IPv6 packets are forwarded along the same path. This isolates the data plane as the main source of (network) performance differences. The first row of Table 2.4 shows nearly identical performance, which established the parity of the IPv4 and IPv6 data planes.

Step 2 considered cases for which IPv4 and IPv6 control plane decisions differ,

i.e., the paths chosen by IPv4 and IPv6 routing are different. Note that such differences arise primarily because of adoption (or lack thereof) decisions. Specifically, instead of following the optimized IPv4 path, IPv6 routing is required to detour (or tunnel) around routing domains (ISPs) that have either not deployed IPv6 or opted not to establish IPv6 peering sessions with their neighbors. Measurement data revealed that a substantial performance gap remained in those cases (second row of Table 2.4). Hence, establishing the control plane, and therefore ISP's adoption decisions, as the main contributor to IPv6 continuing performance lag.

In summary, as of 2011 IPv6 was finally on par with IPv4 *technology-wise,* but while the performance gap had narrowed, it had not disappeared. Limited adoption (among ISPs), which IPv6 initial technical immaturity had contributed to, was still preventing parity by forcing the use of less efficient paths. In other words, IPv6 low adoption among ISPs was potentially slowing its future adoption by perpetuating a performance gap with IPv4. This begged the question of what adoption level was needed to, if not close, at least make this gap less perceptible.

As Fig. 2.4 demonstrates, the performance gap between IPv4 and IPv6 had essentially disappeared by 2013 (they perform identically about 80% of the time, and each outperforms the other for the remaining 20%). The hypothesis is that IPv6 adoption, at least in the core of the Internet, is now sufficient to ensure that even when IPv4 and IPv6 control plane decisions differ, the detours IPv6 may still have to make now have a negligible impact. Tables 2.5 and 2.6 offer data in

support of this conclusion. Table 2.5 shows that after 2011, not only did destinations with identical IPv4 and IPv6 paths continue to see mostly comparable performance (confirming performance parity), an increasing number of destinations accessible over different IPv6 and IPv4 paths also achieved parity. As Table 2.6 suggests, this can be attributed to "shorter detours" taken by IPv6 paths because of the greater density of IPv6 ISPs in the core of the Internet[17]. To further assess the extent to which this was the case, we compared IPv6 (AS) path lengths in 2011 and 2012 and found that 72% of them experienced a decrease. This is in contrast to only 18% of IPv4 paths experiencing a decrease in the same period.

In summary, by 2013 IPv6 had achieved not only technology, but also performance parity with IPv4. The latter was primarily due to higher IPv6 adoption in the core of the Internet. This contributed to decreasing the number and length of IPv6 detours, which all but eliminated differences in latency between IPv6 and IPv4 paths.

	Top 100K Sites	Top 1M Sites
Same Path Destinations	100%	94%
Diff. Paths Destinations	79%	84%

Table 2.5: IPv6 better or equal to IPv4 after 2011.

Another category of websites of potential interest is that of sites associated with

[17]This is consistent with CAIDA's measurements summarized in Table 2.2), which showed an increased density of IPv6 in the core of the Internet.

	IPv4 Transit ASes	IPv6 Transit ASes
2011	216	134
2012	229	147
Growth	6%	10%

Table 2.6: Transit ASes sampled in our measurements.

different destination ASes in IPv6 and IPv4, with Table 2.7 showing how they fared performance-wise. There are various possible reasons for why IPv6 and IPv4 queries for a given webpage are sent to different locations. One of them is clearly the use of CDNs, especially since until 2012 very few CDN providers offered IPv6 services[18]. We were, however, only able to confirm the use of CDNs for a few such websites[19]. Irrespective of the reason behind the difference in destination ASes for IPv6 and IPv4 queries, Table 2.7 shows that IPv6 performance also improved for this category of sites. This is again likely due to the overall improvement in IPv6 connectivity that made IPv6 paths more efficient.

In summary, IPv6 lack of technology maturity initially resulted in poor performance, which likely contributed to slow adoption by ISPs. This in turn ensured a

[18]See http://www.cdn-advisor.com/tag/ipv6/.

[19]Among 100 randomly chosen such websites, only 42 could be directly linked with a well-known CDN service provider such as Akamai, Bitgravity, NTT, Bankinform, Cloudflare, Edgecast, Amazon, Google, Softlayer, Tata, etc. For the remaining 58 sites, we could neither confirm a well-known CDN service, nor could we rule out reliance on a lesser-known CDN provider, or some form of load-balancing mechanism.

	Top 100K Sites	Top 1M Sites
2009–2011	67%	70%
2012–Present	80%	78%

Table 2.7: IPv6 better or equal to IPv4 – Different ASes.

	Phase I	**Phase II**	**Phase III**
Demand for IPv6 Tech.	Moderate	Large	Very Large
IPv4 Address Scarcity	✗	Anticipated	Realized
Infrastructure Upgrade Cost	Large	Moderate	Moderate
Translation Cost	✗	Low	Marginally Increasing
# of IPv6 Users	Negligible	Marginal	Moderate
# of IPv6 Accessible ICPs	Negligible	Marginal	Moderate
Quality of IPv6 Connections	Low	Moderate	High
Migration Status	***Stagnant***	***Emerging***	***Accelerating***

Table 2.8: Evolution of Key IPv6 Adoption Factors.

persisting performance gap, even after IPv6 achieved technology parity. This appears to have changed around early 2012, with IPv6 finally achieving parity with IPv4. This should, hopefully, further facilitate IPv6 continuing adoption.

2.4.3 Closing the loop: positing cause and effect relationships

In this last section, we attempt to correlate the three phases of the IPv6 migration observed in Section 2.3.2 to changes in the different factors identified in the previous section. For convenience, we summarize those changes in Table 2.8. The discussion is followed in Section 2.5 by the introduction of a simple model based on the parameters of Table 2.8. The model seeks to capture the complex dependencies and interactions that exist between those parameters, and their effect on IPv6 adoption. The primary purpose is to *qualitatively* validate the causal relationships posited in this section between IPv6 adoption and changes in these parameters. In other words, given an evolution of the IPv6 ecosystem similar to that of Table 2.8, does the model yield changes in IPv6 adoption consistent with the observations of Section 2.3.2. The goal is not to precisely reproduce those changes, but instead to confirm the cause and effect intuition we articulate next.

During *Phase I* (before 2009), IPv4 addresses were still plentiful and their exhaustion far in the future, so that demand for IPv6 products was low and limited mainly to a few forward-looking ISPs. This ensured a relatively low initial in-

vestment in the development of IPv6 technology by ITDs. This combination of limited investment and few users to test the technology likely contributed to the slow maturation of IPv6 technology. This in turn kept demand low and perpetuated the status quo. There does not appear to have been a single landmark event that triggered a sudden increase in ITDs investment in the development of IPv6 technologies. Instead a slow but steady rise in awareness, in part brought about by various government programs and mandates, *e.g.,* see [16], resulted in IPv6 technologies being progressively brought on par with their IPv4 counterpart. By 2009, most key Internet technologies supported IPv6, and did so at a level of quality and stability close to that of IPv4.

Near technology parity paved the way for the emergence of IPv6 that started in *Phase II*. Technical parity was, however, by itself not sufficient to trigger mass adoption. IPv6 still lacked a strong enough incentive to overcome the adoption cost it imposed on both ISPs and ICPs. This remaining barrier was further strengthened by dependencies between stake-holders (ICPs had little incentive to become IPv6 accessible without a critical mass of IPv6 users, and ISPs where hesitant to invest in assigning IPv6 addresses, when IPv4 addresses were still available and most ICPs were not reachable over IPv6). Hence, in spite of the growing incentive to adopt IPv6 created by the steady decline in free IPv4 addresses and the steady improvements in quality of IPv6, progress remained slow.

Several additional changes were required to usher in the acceleration of IPv6

adoption that started in *Phase III.* The reality of IPv4 address scarcity finally settled in with IANA's allocation of its last block, and a sequence of high-profile events such as World IPv6 Day and World IPv6 Launch further contributed to this realization. In addition, the level of IPv6 adoption in the core of the Internet eventually reached sufficient critical mass to ensure that the quality of IPv6 connectivity was on par with that of IPv4, *i.e.,* did not involve costly detours. As illustrated in Fig. 2.1, this together with the potential for faster growth in the IPv6 user base, made it easier for ICPs to opt to become IPv6 accessible. Anecdotally, this can also be seen when comparing the results of the IPv6 World Day (June 2011) and IPv6 World Launch (June 2012) (see again Fig. 2.1). Many ICPs that "tried" IPv6 during IPv6 World Day reverted to IPv4 after the event, while most IPv6 trials converted to permanent status after IPv6 World Launch.

In the next section, we introduce a simple model that seeks to connect more formally the parameters and patterns identified in Table 2.8, to the three-phase adoption of Section 2.3.

2.5 A Simple Validation

Our goal in this section is to offer a simple validation of the causal relationships put forward in the previous section, between changes in the IPv6 ecosystem and the IPv6 adoption pattern observed in Section 2.3.2. For that purpose, we develop a model that captures interactions between the parameters of Table 2.8 and their

effect on IPv6 adoption. We then vary those parameters in a manner consistent with Table 2.8, and show that the model produces changes in IPv6 adoption that are qualitatively consistent with the trends reported in Section 2.3.2. We note that a more quantitative validation is challenging, as accurately estimating both exact changes in the parameters of Table 2.8 and their relative weights in the model is at best difficult. Furthermore, the specialized nature of the IPv6 adoption problem and its many unique parameters, make it unlikely that a more precise model formulation would have value that extends to other settings and technology adoption scenarios. Instead, the model developed in this section offers a broad confirmation of the cause-and-effect relationships posited in the previous section. As we illustrate later, it also enables coarse "what-if" investigations, which can highlight the importance of certain parameters in keeping IPv6 adoption on track.

2.5.1 Model Overview

The model involves the three major decision makers of Section 2.4, namely, ITDs, ISPs, and ICPs. As alluded to earlier, users are mostly passive, with their "adoption" of IPv6 largely a consequence of decisions made by others. Stake-holder's decisions to adopt IPv6 are based on utility functions that depend on multiple factors, including the adoption decisions of other stake-holders. Stake-holders revisit their decisions at discrete epochs indexed by i, to account for changes in both the Internet ecosystem, *e.g.,* a decrease in the number of available public IPv4 addresses

or growth in the number of Internet users (the Internet user-base is assumed to grow at a rate of q in each epoch), and the decisions of other stake-holders. For simplicity and in keeping with practice, IPv6 adoption decisions are assumed irreversible (once their cost has been incurred, there is little benefit to reverting). ITDs, ISPs and ICPs boast different utility functions, and the model allows for heterogeneity in their individual decisions as well as limited competition (for ITDs).

In the next two sub-sections, we introduce expressions for the utility functions of ITDs, ISPs, and ICPs, and describe their use in making adoption decisions. The last sub-section is devoted to evaluating the model's outcome under combinations of parameters that mimic the progression of Table 2.8.

2.5.2 Utility Functions

ITDs

They provide Internet technologies to other Internet stakeholders, and their (IPv6) utility is expressed through the revenues they generate from their IPv6 products. Revenues depend on demand (*i.e.,* market size), which grows as more of the Internet migrates to IPv6. ITDs periodically assess the IPv6 market size, denoted as $M(i)$ at epoch i. For simplicity, the model does not endogenize the relationship between $M(i)$ and the level of ISP and ICP adoption. Instead, it couples them exogenously in its numerical evaluation.

ITDs are split into different market segments, *e.g.,* router vendors, OS devel-

opers, etc., with segment j assigned a share δ_j of the overall IPv6 market. Within a segment, the model includes two ITDs to incorporate the effect of competition. Market size determines whether an ITD invests in developing IPv6 technology and at what level, with a higher level of investment corresponding to higher product quality. Quality varies between 0 and 1, with 0 denoting no product offering and 1 corresponding to parity with IPv4. The quality of an ITD's offering determines how it shares its market segment with its competitor. A product's quality is taken to be proportional to the ITD's cumulative investment in developing the product, and therefore of the form:

$$Q_j(i) = \sum_{l=1}^{i} c_l \,, \tag{2.5.1}$$

where $Q_j(i)$ denotes the quality of the IPv6 offering of a type j ITD at epoch i, and c_l is the investment it made at epoch l. The model further assumes that at each epoch, type j ITDs play a best response game with their competitor to determine their investment. The utility function of an ITD of type j at epoch i is, therefore, of the form:

$$U_{ITD}(i,j) = \frac{Q_j(i)}{Q_j(i) + Q_j^{Comp.}(i)} \delta_j M(i) - c_i \,, \tag{2.5.2}$$

where $Q_j^{Comp.}(i)$ is the quality of the ITD's competitor(s) at epoch i. Eq. (2.5.2) captures the relationship between the investment an ITD makes and the revenue it generates from investing in its IPv6 products. Note that Eq. (2.5.2) assumes a symmetric decision process by the competing ITDs in market segment j. This is for analytical tractability, but does not qualitatively affect the model's outcome.

Eq. (2.5.2) also reflects two important aspects of IPv6 investments by ITDs: **(i)** they are demand-driven, *i.e.,* if there is no demand ($M(i) \sim 0$), ITDs do not invest in IPv6, and conversely, growth in $M(i)$ fuels investments; and **(ii)** improving the quality of IPv6 products is in part driven by competition (see Section 2.5.3 for details).

ISPs

They all eventually need IPv6 to grow (keep adding new users), but upgrading their network to IPv6 involves a cost. The cost of upgrading an ISP's network depends on network size and a "unit" upgrade cost. The size of an ISP's network is assumed proportional to its user base, and the model allows heterogeneity in ISP sizes. The initial size of the m^{th} ISP is denoted as n_m, and is assumed to grow at a constant rate of q. The unit cost of upgrading an ISP's network to IPv6 depends on the availability and quality of versions of ITDs' technologies. For simplicity, the model assumes that it is inversely proportional to a parameter $\phi(i)$ that tracks availability and quality of IPv6 technologies at epoch i ($\phi(i)$ takes values in $[0, 1]$ –see Eq. (2.5.9)– with 0 corresponding to no IPv6 version of a technology, and 1 to quality that is on par with that of IPv4). The need to acquire more IPv4 addresses is the main counterpart to the cost of upgrading one's network, and the model assumes that ISPs are heterogeneous in the number of IPv4 addresses they initially have at their disposal. This number is denoted as k_m for ISP m at epoch 0[20]. The

[20]In Section 2.5.4, both n_m and k_m are taken to be uniform in $[0, 1]$.

unit cost of new IPv4 addresses is chosen[21] to be "quadratic" to reflect the fact that as they grow scarce, their price is likely to increase. The cumulative cost of deferring upgrading to IPv6 until epoch i for the m^{th} ISP is, therefore, of the form:

$$C_m^{Up}(i) = \overbrace{n_m(1+iq)}^{\text{user base}} \quad \overbrace{\frac{1}{\phi(i)}}^{\text{per user upgrade cost}} + \overbrace{C_4((i-1)n_m q - k_m)_+^2}^{\text{IPv4 address acquisition cost}}. \tag{2.5.3}$$

The first term is the cost of upgrading an infrastructure that has grown to a size of $n_m(1+iq)$ by epoch i given a unit upgrade cost of $1/\phi(i)$, while $C_4((i-1)n_m q - k_m)_+^2 = C_4 \max(0, ((i-1)n_m q - k_m))^2$ denotes the cost of acquiring IPv4 addresses up to epoch $i-1$ (this cost remains 0 until the ISP exhausts its initial IPv4 address pool of size k_m). Note that $n_m(1+iq)$ grows over time, while $1/\phi(i)$ decreases as IPv6 technology improves. As we shall see in Section 2.5.3, ISPs seek to identify the epoch that minimizes upgrade costs.

Once an ISP has upgraded its network to IPv6, the model assumes that it allocates *both* IPv6 and IPv4 addresses to new users until it runs out of the latter. Once this happens, it must either purchase more IPv4 addresses, or deploy translation devices to enable IPv6-only users to access the IPv4 Internet. The choice is based on cost, and given that, as discussed in Section 2.4.2, translation costs are expected to decline, we further simplify the model by assuming that translation is the solution of choice. The main impact of this assumption is in increasing the

[21] Other cost functions can be chosen, *e.g.,* constant, and while they quantitatively affect the results, the outcome remains qualitatively similar.

number of IPv6-only users, which, as we shall see next, positively influences ICPs' decisions.

ICPs

Their revenue depends in part on the quality with which they deliver content to users. ICPs that are not IPv6 accessible must rely on translation to access IPv6 users, and as the IPv6 user base grows, the connectivity impairment this imposes on those users (see again [3, 14, 27, 53]) translates into an increasing penalty (revenue loss). ICPs, therefore, weigh this loss against the cost of becoming IPv6 accessible. This cost is primarily an infrastructure upgrade cost, similar in nature to that of ISPs. It depends on the availability of IPv6 technologies and the size of the ICP's infrastructure. An ICP decides to become IPv6 accessible once the cost of doing so is lower than the revenue gain the change will generate. This decision process is captured in Eq. (2.5.4) that also incorporates heterogeneity among ICPs based on their popularity. More popular ICPs are assumed to generate higher revenues from their users, as well as incur lower upgrade costs (because of economies of scale). As expected, this translates into more popular ICPs adopting earlier, consistent with Fig. 2.2.

$$\begin{aligned}\Delta_{ICP}(i) =& N_6(i)\beta a_6(i) \\ & -[N_{46}(i)\beta\alpha(i)) + S_{infra}(i) \quad \underbrace{(2-\beta)\frac{1}{\phi(i)}}_{\text{per user upgrade cost}} \quad]\end{aligned} \tag{2.5.4}$$

$\Delta_{ICP}(i)$ measures the impact on the ICP's revenue of becoming IPv6 accessible at epoch i. The first term in Eq. (2.5.4) represents the gain associated with IPv6 accessibility, with $N_6(i)$ the size of the IPv6 user base at epoch i, β the ICP's popularity factor[22], and $a_6(i)$ the per user revenue gain from native IPv6 connectivity at epoch i (it increases over time as the quality of IPv6 technology improves). Note that Eq. (2.5.4) highlights that when native IPv6 connectivity is (quality-wise) worse than what is achievable through translation devices, *i.e.,* $a_6(i) \leq 0$, ICPs have little to no incentives to become IPv6 accessible (because $\Delta_{ICP}(i) \leq 0$). The second term in Eq. (2.5.4) includes both a potential revenue loss associated with becoming IPv6 accessible, and the cost of upgrading the ICP's infrastructure to IPv6.

The potential revenue loss associated with IPv6 accessibility is in the term $N_{46}(i)\beta\alpha(i)$. It accounts for the fact that dual-stack users (there are $N_{46}(i)$ of them) often access IPv6 accessible ICPs over IPv6 and not IPv4, which may result in lower connectivity quality. This is captured through $\alpha(i)$ that denotes the per user revenue loss at epoch i from IPv6 connectivity relative to IPv4 connectivity. Finally, the term $S_{infra}(i)(2-\beta)/\phi(i)$ represents the ICP's IPv6 upgrade cost that is proportional to the size of its infrastructure at epoch i, $S_{infra}(i)$, and, as with ISPs, is inversely proportional to the availability and quality of IPv6 technology as measured through $\phi(i)$. The factor $(2-\beta)$ reflects the economies of scale assumed available to more popular ICPs (the least popular ICPs have upgrade costs twice

[22] β is distributed in $[0, 1]$, with 1 the highest popularity.

those of more popular ICPs).

An ICP re-evaluates the benefit of IPv6 accessibility at each epoch to account for changes in the parameters of Eq. (2.5.4). Factors that contribute to making IPv6 more attractive include growth in $N_6(i)$, the number of IPv6-only users and improvements in IPv6 quality that contribute to both increasing $a_6(i)$, the revenue gain afforded by native connectivity for IPv6-only users, and decreasing the revenue loss $\alpha(i)$ incurred for dual-stack users. On the other hand, the ICP's infrastructure size, $S_{infra}(i)$, keeps growing, so that upgrade costs may increase, unless the per user cost of upgrading to IPv6, $(2-\beta)/\phi(i)$, decreases commensurately. To assess the overall impact of these different factors, the model uses the following expressions for estimating changes in $a_6(i)$ and $\alpha(i)$:

$$a_6(i) \sim \phi(i)\mu(i) + I_{CDN} \tag{2.5.5}$$

$$\alpha(i) \sim 1 - \frac{\phi(i)\mu(i) + I_{CDN}}{2}\,, \tag{2.5.6}$$

where as in Eq. (2.5.3), $\phi(i)$ measures the availability and quality of IPv6 technology, I_{CDN} denotes the fraction of CDN providers that support IPv6 (they can have a strong impact on IPv6 quality), and the product $\phi(i)\mu(i)$ captures the dual impact of the network data plane ($\phi(i)$) and control plane ($\mu(i)$) on the overall quality of IPv6. As discussed earlier, $\phi(i)$ depends on the maturity of IPv6 technology, while $\mu(i)$ increases as more ISPs adopt IPv6 (detours around IPv4 only islands become shorter). In the next section, we formalize the evolution of those parameters and their dependencies.

2.5.3 Decision Mechanisms & Solution Method

This section reviews the decision process that the utility functions of the previous section give rise to under the assumption that stake-holders make decisions that maximize their utility. In other words, they are rational.

ITDs

ITDs' decisions are when and how much to invest in developing IPv6 versions of their technology. We assume that to be viable IPv6 products must meet a minimum quality threshold $0 < Q_{min} < 1$. Hence, an ITD of type j first invests in IPv6 at epoch i if the cost (of meeting the minimum quality threshold) is less than the revenue potential of the IPv6 market, as defined in Eq. (2.5.2). The decision, therefore, depends on the ITD's type, δ_j, the estimated size of the IPv6 market at epoch i, $M(i)$ (an increasing quantity), and the quality of its competitor's technology at epoch i, $Q_j^{Comp.}(i)$. The first two parameters are exogenous, while the ITD needs to anticipate $Q_j^{Comp.}(i)$. Given the assumption of a symmetric decision process (more on this below), the two competing ITDs (in market segment j) make consistent decisions, *i.e.,* they invest to offer products of comparable quality so that $Q_j(i) = Q_j^{Comp.}(i)$.

This competition between ITDs can be modeled as a best response game at each epoch. The actions of both players are their level of investment in IPv6, which in turns determines the quality of their offering. Both players in segment j account

for the decision process of their competitor, so that their best response decisions are in the form of an investment that at epoch i yields a cumulative quality $Q_j(i)$ for their technology of the form

$$Q_j(i) = \sqrt{Q_j^{Comp.}(i)M(i)\delta_j} - Q_j^{Comp.}(i). \tag{2.5.7}$$

The symmetric nature of the two ITD competitors in market segment j produces a Nash equilibrium where they split the market equally with a cumulative quality $Q_j(i)$ of the form:

$$Q_j(i) = \max\left\{\frac{M(i)\delta_j}{4}, 1\right\}, \tag{2.5.8}$$

Note that Eq. (2.5.8) implies that ITDs wont invest in IPv6 versions of their technologies until $\delta_j M(i) > 2Q_{min}$, *i.e.,* the IPv6 market size exceeds a certain threshold. Conversely, as $M(i)$ grows, ITDs' technologies investment ultimately results in (quality) parity between the IPv4 and IPv6 versions of their technologies. This in turn yields the following expression for the parameter $\phi(i)$ that measures the overall availability and quality of IPv6 technologies at epoch i:

$$\phi(i) = \sum_j Q_j(i)\delta_j = \frac{M(i)}{4}\sum_j \delta_j^2\,, \tag{2.5.9}$$

where the summation is over all market segments.

ISPs

An ISP's goal is to find the epoch at which the cumulative cost of upgrading its network to IPv6 is "minimal." Upgrade costs are initially high because IPv6 quality,

$\phi(i)$, is low. As per Eq. (2.5.3), this leads some ISPs to defer upgrading until quality improves[23]. As IPv6 quality improves and approaches parity with IPv4, upgrade costs eventually increase driven by growth in an ISP's user base. Predicting the exact crossover point is complex, and our goal is not to offer precise guidelines. Instead we seek to capture the inherent tension between those two factors in an ISP's decision. For that purpose, we assume that ISPs rely on a myopic decision process and simply evaluate whether the rate of increase of upgrade costs is higher than in the previous period, and upgrade as soon as it is.

In other words, the m^{th} ISP adopts IPv6 at epoch i_m if $C_m^{Up}(i_m) - C_m^{Up}(i_m - 1) > C_m^{Up}(i_m - 1) - C_m^{Up}(i_m - 2)$, where $C_m^{Up}(i)$ is as per Eq. (2.5.3). With ISPs decisions known, the fraction $\mu(i)$ of ISPs that have upgraded to IPv6 by epoch i can then be readily obtained, and therefore used to determine its impact on IPv6 connectivity quality as per Eqs. (2.5.5) and (2.5.6). Recall that the latter play a role in ICPs decisions, and capturing those interactions is one of the model's goals.

Once an ISP has upgraded its network to IPv6, it faces another decision, namely, how to continue to provide new users with access to the IPv4 Internet. This is an easy decision as long as the ISP still has IPv4 addresses, *i.e.,* until epoch $i = k_m/n_m q$ for the m^{th} ISP, as new users can be assigned *both* IPv4 and IPv6 addresses. Once an ISP runs out of IPv4 addresses, it must then decide between acquiring more IPv4 addresses and deploying translation mechanisms, as discussed in Section 2.4.1.

[23]Heterogeneity in decisions arises from differences in both ISPs' size and in the number of IPv4 addresses they own.

The model can be readily adapted to allow for such a decision, *i.e.,* select the lowest cost option. However, we assume in our evaluation (Section 2.5.4) that ISPs are "strategic" and opt to handle IPv4 connectivity (for new IPv6 users) solely through translation mechanisms. The primary motivation is, as outlined next, that this offers additional incentives for ICPs to become IPv6 accessible earlier. Hence, hastening the Internet's migration to IPv6, and ultimately lowering ISPs costs.

ICPs

An ICP's goal is to maximize the revenue it derives from having Internet users. For that purpose, it re-evaluates $\Delta_{ICP}(i)$ (Eq. (2.5.4)) at every epoch, and becomes IPv6 accessible at the first epoch i for which $\Delta_{ICP}(i) > 0$. Under the assumption that ICPs' popularity β is uniformly distributed in $[0, 1]$, this yields the following expression for the fraction $\gamma_6(i)$ of ICPs that are IPv6 accessible at epoch i:

$$\gamma_6(i) = \frac{N_6(i)a_6(i) - N_{46}(i)\alpha(i) - S_{infra}(i)/\phi(i)}{N_6(i)a_6(i) - N_{46}(i)\alpha(i) + S_{infra}(i)/\phi(i)} \tag{2.5.10}$$

Using Eq. (2.5.10), it is easy to establish the following intuitive statements that highlight the dependencies that exist between ICPs decisions and those of ISPs and ITDs:

The fraction of ICPs natively accessible over IPv6 increases as either the number of IPv6 users increases, or the quality of IPv6 increases ($a_6(i)$ increases, $\alpha(i)$ decreases). In addition, once the IPv6 user base is large enough ($N_6(i)a_6(i) > N_{46}(i)\alpha(i)$), decreases in upgrade costs contribute to increasing the number of IPv6

accessible ICPs. Conversely, increases in the number of dual-stack users can delay increases in the number of IPv6 accessible ICPs.

2.5.4 Model's Evaluation

The goal of this section is to explore the extent to which the progression of IPv6 adoption documented in Section 2.3 can be reproduced using the arguably stylized model that was just presented. For that purpose, we consider "configurations" associated with different combinations of the model's exogenous parameters, and characterize the evolution of IPv6 "adoption" across stake-holders as the Internet's user base increases. Specifically, we numerically evaluate the model's outcome for three different sets of exogenous parameters that mimic the three right columns of Table 2.8.

The first configuration emulates IPv6 early years. Demand for IPv6 versions of Internet technologies was initially non-existent ($M(i) \sim 0$). As a result, development incentives were low even in segments with large market shares δ_j, *e.g.,* router and OS vendors. The outcome predicted by Eq. (2.5.9) is marginal availability of IPv6 technologies, *i.e.,* $\phi(i) \sim 0$, and consequently large upgrade costs ($1/\phi(i) \gg 0$). This in turn translates into a negligible fraction of ISPs upgrading their network to IPv6 ($\mu(i) \sim 0$) and similarly very few ICPs opting to become IPv6 accessible ($\gamma_6(i) \sim 0$). Initiatives aimed at promoting support for IPv6, *e.g.,* government mandates, helped change the situation and create some early demand for

IPv6 technologies ($M(i) > 0$) even in the absence of a real driver (the exhaustion of IPv4 addresses was still far away). This in turn triggered some initial investments on the part of ITDs (see again Eq. (2.5.9)), so that early releases of IPv6 products became available, *i.e.,* $\phi(i) > 0$. This lowered upgrade costs ($1/\phi(i)$), but ultimately had little effect on IPv6 adoption by either ISPs or ICPs, *i.e.,* $\mu(i) \gtrsim 0$, and $\gamma_6(i) \gtrsim 0$. The reason, consistent with Eqs. (2.5.3) and (2.5.4), is that while demand for and availability of IPv6 technology improved, IPv6 quality/stability remained below that of IPv4 ($a_6(i)$ was still small), endemic problems continued to plague dual-stack users ($\alpha(i)$ stayed large), and IPv4 address exhaustion was nowhere near.

The second configuration seeks to capture the second phase of IPv6 adoption in Section 2.3. During that phase, demand for IPv6 products increased to a point where most of the ITDs supported IPv6 in their products at a level of stability/quality on par with that of IPv4, *i.e.,* $\phi(i) \sim 1$. This was sufficient to incentivize some ISPs to adopt IPv6. Most of those ISPs, however, still owned IPv4 addresses, so that new users were primarily dual-stack (as opposed to IPv6 only), *i.e.,* $N_6(i)$ stayed small while $N_{46}(i)$ grew. This offered little motivation for ICPs to consider becoming IPv6 accessible, especially since IPv6 connectivity quality was still lagging behind IPv4 (because many ISPs had not yet upgraded to IPv6). This is consistent with Eq. (2.5.10) that produces only small increases in $\gamma_6(i)$ under those configurations.

The third configuration maps to phase three of Section 2.3. ISPs are increasingly running out of IPv4 addresses, and because IPv6 technology is stable and on par with IPv4, upgrading to IPv6 now makes sense for many of them. The larger number of IPv6 ISPs together with the greater availability of IPv6 versions of services such as CDNs result in IPv6 connectivity quality being now equals that of IPv4 ($a_6(i) \sim 1$ and $\alpha(i) \sim 0$). This eliminates the quality penalty that IPv6 users suffer compared to IPv4 users. When combined with a growing number of IPv6-only users ($N_6(i)$), this is enough to entice an increasingly large number of ICPs to become IPv6 accessible; a phenomenon that Eq. (2.5.10) again captures.

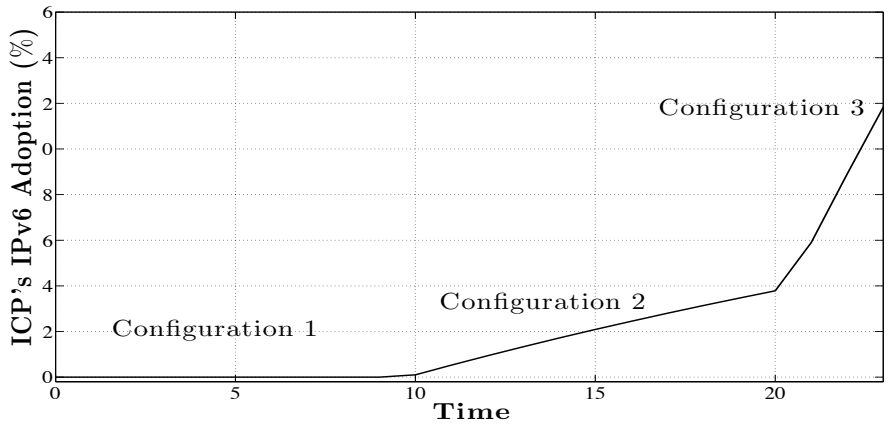


Figure 2.5: Model-driven evolution of ICPs' IPv6 adoption.

IPv6 adoption under those three configurations is shown in Fig. 2.5 for ICPs. The outcome is qualitatively similar to Fig. 2.1. This is obviously no "proof" of the model's validity. However, it offers a level of validation for the causal relationships put forward in Section 4.2, connecting changes in the IPv6 ecosystem and the observed evolution of IPv6 adoption.

The model can also be used for coarse "what-if" analyses exploring the potential impact of changes in the IPv6 ecosystem. Those can prove useful to avoid missteps, which could, if not derail, at least slow-down IPv6 adoption and in the process increase its overall cost. We illustrate this through a simple example that considers a scenario where ISPs that migrated to IPv6 proceed to sell their IPv4 addresses on open markets such as those of Section 2.4.2. The resulting influx of new IPv4 addresses would likely stabilize or even reduce IPv4 address costs. This would in turn make it easier for ISPs that have not yet migrated to IPv6 to defer this decision; in the process slowing down the growth of the IPv6 user base. The impact on ICPs is less clear since while there would be fewer IPv6 users overall, more of those users would now be IPv6-only. The former is a disincentive to becoming IPv6 accessible, while the latter acts as an incentive.

The model offers the opportunity to investigate the impact of such a change, with the results shown in Fig. 2.6. The figure demonstrates that the slower growth in the total number of IPv6 users produced by lower acquisition costs for IPv4 addresses is the dominant factor. It results in ICPs delaying their decision to become IPv6 accessible. The insight that emerges from this "what-if" scenario is that although ISPs that migrated to IPv6 stand to derive short-term benefits from selling their IPv4 addresses, those benefits are likely to be offset by the higher cost they will incur from the Internet's slower migration to IPv6, *e.g.,* through higher translation costs.

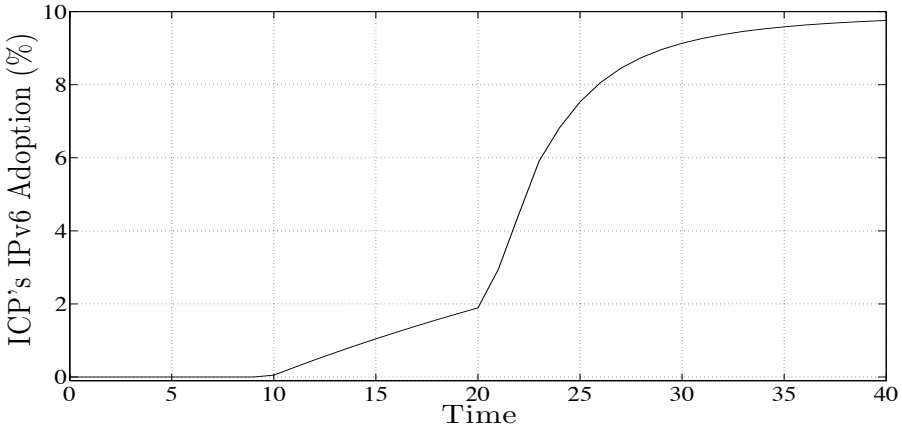


Figure 2.6: Impact of lower IPv4 address acquisition costs when ISPs sell their IPv4 addresses after migrating to IPv6.

2.6 Conclusion

This work in this chapter reports on measurements capturing the evolution of IPv6 adoption across Internet stakeholders, and identifies factors likely to have influenced adoption decisions. It posits, and to some extent documents, changes in those factors as possible causes for transitions in IPv6 adoption patterns observed in the measurements. We also develop a simple model connecting those changes to their impact on IPv6 adoption, and use it to qualitatively validate its hypotheses. The investigation identifies the coupling between low initial demand for IPv6 products and their lower quality compared to their IPv4 counterpart as an important contributor to IPv6 early adoption challenges. In particular, it appears largely responsible for the initial reluctance of service and content providers to adopt IPv6, which in turn deterred users and contributed to prolonging the Internet's migration to IPv6.

Chapter 3

Maintaining the Progress of IPv6 Adoption

3.1 Introduction

IANA announced in February 2011 that the free pool of IPv4 addresses is depleted, and even if IPv4 addresses scarcity has not yet materialized everywhere, we are slowly but surely headed in that direction. IPv6 was designed to address this issue, and even though our study presented in Chapter 2 shows its adoption is accelerating, there are hurdles that can impede or slow down its progress in the future. Although these hurdles are not (anymore) of a technical nature, years of technology disparity between IPv4 and IPv6 caused a marginal adoption of IPv6 across major Internet stakeholders [62], which in addition to incompatibility of the two technologies forced the use of translation mechanisms to allow IPv6-only users access to the IPv4-only Internet. These translation mechanisms are widely used today by ISPs such as CERNET2 in China, and Verizon Wireless and T-Mobile in the U.S. CERNET2 [90] (an academic network), already had over 400k *IPv6-only* users in 2009, is expected to reach 3 million by the end of 2015 (see [12, 13]), and uses "IVI", which translates IPv4 traffic to IPv6 and vice versa. Similarly, Verizon Wireless and T-Mobile are now primarily relying on IPv6 addresses for new cell-phone subscribers [77, 82], and use "NAT444" and "464XLAT" as their translation mechanisms, respectively. While necessary for a transition, the quality degradation those mechanisms introduce [3, 14, 27, 53] reduces motivation for the new users to adopt IPv6. This is an instance of hurdles in front of the progression of IPv6 adoption in the future. Our initial intuition was that besides the above

instance, the distributed structure of the Internet can also affect the progression of IPv6 adoption. Specifically, the benefit of migrating to IPv6 depends to a large extent on what others in the Internet do. This is not an uncommon situation (*e.g.,* see [2] for a related discussion in the context of Internet security protocols), but uncertainty in the decisions of others can significantly delay the adoption of a new technology.

A goal of this chapter is, therefore, to explore and explain strategies that can derail or speed up the current progress of IPv6 adoption. These strategies require careful assessments as we are dealing with a highly decentralized system (the Internet). To better understand the extent to which these strategies can affect IPv6 adoption, several simple yet representative scenarios and models were developed. The focus of these models is on the decision making process of independent and decentralized stakeholders across the Internet, and how those decisions can affect IPv6 adoption. We acknowledge up-front the many simplifying assumptions these models rely on (a necessity in most modeling efforts), and their lack of completeness. However, they incorporate major aspects of IPv6 adoption decisions, namely, (i) heterogeneity in the Internet stakeholders making decisions; (ii) a representative sample of available technology options; and (iii) the dependencies that exist across decisions.

Our findings from these models indicate that independent decision making process of ISPs can negatively affect IPv6 adoption. In other words, disagreement

between ISPs on connectivity option offerings, adds uncertainty to the factors that affect IPv6 adoption decisions of the Internet stakeholders, and makes it hard to identify winning strategies. As a result of this uncertainty, migration to IPv6 slows down, or at the very least becomes haphazard. Another finding of the models is that even minimal coordination among ISPs in offering connectivity options, *e.g.,* an Internet-wide consensus on offering IPv6 as one of the connectivity options, can significantly improve our abilities to identify strategies that hasten the IPv6 migration process. Although consensus alone is not sufficient, it makes it easier for the Internet stakeholders to identify winning strategies that can, at the same time, speed up the migration of the Internet to IPv6.

The chapter's contributions are, therefore, two-fold:

(i) It shows how distributed decision making of the Internet stakeholders, in the presence of competing solutions to the problem of IPv4 address scarcity, can negatively affect identifying winning strategies, and therefore, contribute to the lingering of the current quandary in IPv6 adoption; and

(ii) It illustrates how the introduction of limited coordination among ISPs, which is not in itself enough for IPv6 success, can help determine the impact of different parameters on IPv6 adoption, and hence, facilitate a smoother migration process.

The rest of the chapter is structured as follows. Section 3.2 discusses the framework of the problem, including the Internet stakeholders, connectivity options, and scenarios. Sections 3.3 introduces the models in two categories of disagreement and

consensus. Section 3.4 and 4.5 explore the outcome of the models with a certain set of assumptions, and provide the key findings. Section 3.6 investigates the robustness of our findings to different modeling assumptions and extensions. Section 3.7 briefly reviews related works, with Section 4.7 summarizing the chapter's findings and recommendations.

3.2 Problem Framework

There are many factors that arguably affect the adoption of IPv6, and any (tractable) model is unlikely to account for all of them and their variations across stakeholders. Our models operate within a certain framework, and this section specifies the outline of that framework by introducing the Internet stakeholders, their connectivity options, the inter-dependencies between their decisions, and the scenarios in which they interact.

3.2.1 Internet stakeholders

We distinguish between *three* types of Internet stake-holders: Internet Service Providers (ISPs), Internet Content Providers (ICPs), and Internet Content Consumers (users). ISPs derive revenues from providing Internet connectivity to both ICPs and users, and are, therefore, concerned with the choices and costs of the technologies used to implement this connectivity. They make the ultimate decision to offer IPv6 connectivity to the other two stakeholders, hence, they play the most

significant role in IPv6 adoption across the Internet. ICPs obtain the bulk of their revenues from users that connect to them through ISPs. Hence, their focus is on the quality of their connectivity to users and how it may affect their revenues, as well as any cost they may incur to upgrade their existing infrastructure to support a new connectivity option, *e.g.,* IPv6. Finally, users purchase Internet connectivity from ISPs, and use it primarily to connect to ICPs (and to a lesser extent to each others). Hence, they are affected by the cost of Internet connectivity and by its quality.

3.2.2 ISP's connectivity options

ISPs are the providers of Internet connectivity, and therefore control technology choices. Although IPv6 adoption is on the verge of happening, implicit to our modeling effort lies the fact that IPv6 still faces competing solutions. Among those available technology choices ISPs may choose from to accommodate customer growth, we consider three representatives.

The first choice an ISP can make is to simply continue using public IPv4 addresses. This has the advantage of full compatibility with the current Internet, but given the growing scarcity of public IPv4 addresses is likely to quickly involve added costs, *e.g.,* to purchase public IPv4 addresses from an address market such as Hilco Streambank IPv4 Address Marketplace.

The second option an ISP can rely on is to use private IPv4 addresses together

with Carrier-Grade NATs (CGNs). Unlike public IPv4 addresses, private IPv4 addresses can be reused and so are not scarce. CGNs are required to allow connectivity to the public Internet, but the technology behind CGNs is mature. Private IPv4 addresses also have the benefit of letting ISPs defer a potentially expensive upgrade of their network to IPv6. The main disadvantage (to the ISP) is the cost of CGNs, which grows as more users are assigned private IPv4 addresses.

IPv6 is the third option. IPv6 addresses are not scarce, but like private IPv4 addresses will require some form of "translation," *e.g.,* NAT64 [5] or DSLite [29], to allow IPv6 users to communicate with the IPv4 Internet. IPv6↔IPv4 translation is less mature than that for private IPv4 addresses, and may therefore be initially more expensive. On the flip side, even if the exact time-frame remains unclear, the need for translation, and therefore its cost, should disappear as the Internet eventually migrates to IPv6.

3.2.3 Decision dependencies

As alluded to, although ISPs choose Internet technologies, their decisions, including *pricing*, depend heavily on users and ICPs. For example, an ISP offering both IPv6 and (public) IPv4 connectivity might discount the IPv6 service, thereby attracting users to that option and lowering the need for (expensive) public IPv4 addresses. However, more IPv6 users also means higher translation costs, unless this entices more ICPs to become IPv6 accessible thereby lessening the need for translation.

This creates a complex web of dependencies, whose impact is amplified by the distributed decision process that prevails in the Internet. As we shall see, this can make devising sound (profit maximizing) strategies difficult if not impossible. We show in the next sections that these dependencies indeed play a critical role in IPv6 adoption, and by breaking only one of the links in the web of dependencies, the outcomes change drastically.

3.2.4 Scenarios

In many technology adoption instances presence of multiple entrants, and lack of consensus on a single choice among stakeholders can prevent a full market penetration by any of those choices. While competition of alternative solutions can be helpful in keeping the evolution of a technology on the right track, consensus on one choice makes a full market penetration faster and easier. In the case of IPv6, a full market penetration is required, if the Internet is to avoid permanent traffic translation, therefore, the Internet Engineering Task Force (IETF) standardized IPv6 as the replacement for IPv4. However, due to the hurdles in front of IPv6 adoption, other alternative solutions have become popular among some ISPs.

As different ISPs manage separate Autonomous Systems (ASes), their decisions are to some extent independent of each other. This heterogeneity among ISPs can lead them to offer (at least temporarily) different connectivity solutions. Since ISPs provide Internet connectivity, their heterogeneous decision making has a more sig-

nificant impact on IPv6 adoption compared to other Internet stakeholders. Therefore and in order to investigate this impact, we consider two major scenarios: (i) a scenario in which ISPs disagree on immediately offering IPv6 connectivity to their users; and (ii) a scenario in which all ISPs offer IPv6 along with other connectivity options to their users. Next, we describe these two scenarios in more details.

Disagreement on offering IPv6

In this scenario, one ISP is always assumed to offer IPv6, as otherwise the outcome is trivial, *i.e.,* stagnation in IPv6 adoption, while the other ISP offers either public or private IPv4 addresses.

Given that the main competition IPv6 faces is the incumbent IPv4 Internet, we consider the case of two ISPs, one having embraced IPv6 as the technology of choice for its new customers[24], while the other has decided to defer any migration and to simply acquire additional public IPv4 addresses to accommodate new customers. The first ISP needs to deploy address translation devices to allow its new (IPv6) customers to connect to the legacy IPv4 Internet. This cost grows with the number of users that choose IPv6, and decreases as more ICPs become IPv6 accessible[25]. Conversely, while the second ISP does not incur translation costs, it needs to purchase public IPv4 addresses for its new customers. Those costs are expected to rise

[24]T-Mobile has recently started to only assign IPv6 addresses to its Android 4.4 users (see [77]).

[25]Translation costs are assumed proportional to the volume of traffic that needs to be translated, *i.e.,* higher capacity devices are needed.

as public IPv4 addresses become scarcer.

Another variation of this scenario is when no ISP wants to incur the cost of purchasing more public IPv4 addresses (or those addresses are unavailable for purchase). ISPs that defer upgrading to IPv6 would then rely on private IPv4 addresses. Offerings based on either IPv6 or private IPv4 addresses both require translation (CGNs) to connect to the public IPv4 Internet. Translation costs for private IPv4 are likely to be lower than for IPv6, if only because of more mature technology and/or greater operational familiarity and compatibility with the current Internet. On the flip side, translation costs for private IPv4 keep increasing as more users join, independent of how many ICPs become IPv6 accessible. We describe this scenario in Appendix A.

Consensus on Offering IPv6

In this scenario, there exists a global consensus on offering IPv6 (along with other service types), as a technology of choice to users, hence, all ISPs offer IPv6 and another service, *e.g.,* public IPv4.

On the technology choice front, this scenario is identical to the first one, namely, both IPv6 and public IPv4 are available as connectivity options. The main difference is that the two options are now systematically offered by all ISPs, and therefore priced internally, as opposed to competitively, to maximize their own profit. The price difference is a means of modeling, and can be interpreted as the cost of extra

services that ISPs offer (for free) along with their IPv6 services, but charge users for those same services in IPv4, *e.g.,* static addresses (http://www.vo.lu) etc. This scenario is equivalent to having a monopolistic ISP that sets the price of both connectivity choices.

3.3 Models

Based on the scenarios of the last section, we developed models that capture the interactions and decision dependencies of ISPs, ICPs and users. As alluded to in section 3.2.3, the decisions of users depends on the decisions of ICPs and ISPs, and vice versa. ISPs are the selectors of the technology and affect the interactions of the other two stakeholders through their decisions. This framework is common to other environments, *e.g.,* gaming platforms, where the number of game developers and the number of gamers are affected by the decisions of the console provider. Analyzing these frameworks is typically through a two-sided market setting [69]. The ISP is the market maker through its offering of connectivity options, while users and ICPs are the two sides of the market that derive value from each other through the ISP.

We assume that at each step, new and existing users evaluate the Internet connectivity choices available to them through their local ISP(s)[26] and select the

[26]According to http://www.broadbandmap.gov/summarize/nationwide, over 99% of the U.S. population can choose from two or more ISPs, while this figure is 90% in Europe (see

one yielding to the highest *utility*. One obvious shortcoming of this model is the lack of inertia in decision making of users, *i.e.,* every user decides at each time epoch, therefore, in section 3.6 we investigate the robustness of our results in scenarios where the users face some form of inertia, *e.g.,* contractual agreements. We define a user's utility in Section 3.3.1, but it depends primarily on the cost and quality of her Internet connectivity.

Users are assumed heterogeneous, but primarily in their sensitivity to connectivity quality[27]. We further assume (see [27] for a related discussion) that address translation devices, if used, are the main contributors to degradation in connectivity quality/functionality.

Because ICPs are part of the current Internet, they already have a public IPv4 address, and their only decision is whether or not to become IPv6 accessible. They incur a cost when doing so (upgrading their existing IPv4 infrastructure and/or update of operational processes), but unlike users that can revert their decisions, an ICP's decision to become IPv6 accessible is irreversible (once incurring the upgrade cost). Next, we present the utility functions of the Internet stakeholders.

http://goo.gl/MjTPJ6).

[27]Coarser grain heterogeneity is also possible, *e.g.,* between, say, residential and enterprise users, but adds significant complexity to the model. Similarly, heterogeneity in price sensitivity can also be included, but with again a cost in terms of complexity.

3.3.1 Users utility

Users derive a *unit* value from Internet connectivity, with price and quality affecting their overall utility. An alternative model assumes heterogeneous values for different connectivity options, however, since we use pricing as the control knob of the ISPs, the former presentation is chosen (the outcomes are nevertheless similar). Recall that quality is assumed to be primarily affected by (the presence of) translation devices. A user's utility is then captured through the following expression:

$$U_{\text{user}}(\sigma) = \underbrace{1 - p_R}_{V_R} - \sigma a_R \gamma_R \,, \tag{3.3.1}$$

where R indexes connectivity options, p_R is the price of type R connectivity ($p_{\text{pub. IPv4}} > p_{\text{IPv6}} > p_{\text{priv. IPv4}}$) (alternatively V_R is the value of option R), $a_R \in [0, 1]$ quantifies quality (translation) impairments for connectivity option R, if any (a_R is 0 for public IPv4 and positive for both private IPv4 and IPv6), γ_R is the fraction of the Internet (ICPs) affected by those impairments, and σ denotes a user sensitivity to quality impairments.

3.3.2 ICPs utility

ICPs derive revenues from users, and those revenues can be affected by connectivity quality [73]. A major factor in an ICP's decision to become IPv6 accessible[28] is,

[28] As participation in events such "World IPv6 Launch Day" demonstrates, there are obviously many other possible reasons for an ICP to become IPv6 accessible. However, even when those other motivations prevail, the importance of preserving connectivity quality remains, *e.g.,* through

therefore, the impact this decision can have on the revenue it generates from IPv6 users, and how this compares to the cost of upgrading to IPv6 (or convincing its hosting provider to upgrade). Revenue improvements depend on the number of IPv6 users and how they are affected by the ICP's adoption of IPv6. In particular, and as shown in [62], IPv6 and IPv4 connectivity quality are now mostly on par, so that the main benefit of native IPv6 access is to eliminate the need for translation.

The cost of upgrading to IPv6 is largely a function of the "size" of the ICP's infrastructure. For simplicity, this size is assumed proportional to the Internet user-base (the traffic volume an ICP sees grows with the Internet). The net utility in(de)crease an ICP derives from becoming IPv6 accessible can, therefore, be captured as follows:

$$\Delta_6(\text{ICP}) = \beta n_6 a_6 - S_{\text{infra}} \theta c_6 \tag{3.3.2}$$

where βn_6 is the fraction of IPv6 users that an ICP can benefit from, a_6 is the per-user revenue gain from eliminating translation, and θc_6 is the per-user upgrade cost of the ICP's infrastructure (of size S_{infra}). β and θ capture heterogeneity in revenue and cost, respectively, across ICPs.

---

the enforcement of some form of "white-listing."

3.3.3 ISP utility

An ISP's utility (profit) depends on revenues derived from users[29] and costs. Given our aim of assessing the impact of offering different connectivity options, we focus on their cost contributions and ignore other cost components. As costs differ across connectivity options, we introduce the ISP's utility function separately for each.

Public IPv4 only

An ISP that only offers public IPv4 connectivity has a utility function of the form:

$$\Pi_{\text{pub. 4}} = n_4 p_4 - C(n_4 - 1)_+^2 \tag{3.3.3}$$

n_4 is the number of users willing to pay p_4 for public IPv4 connectivity, while $C(n_4 - 1)_+^2 = C \max(0, n_4 - 1)^2$ is the acquisition cost of the $(n_4 - 1)$ additional public IPv4 addresses the ISP needs beyond the "unit" block it already owns (to accommodate its existing users). The quadratic function used for address acquisition costs seeks to capture the growth in the price of public IPv4 addresses due to increasing scarcity. Section 3.6 changes this assumption, and investigates the impact of other functions on the models outcome.

[29] We ignore revenues from ICPs, as they are mostly independent from an ISP's connectivity choices.

IPv6 only (and IPv6↔IPv4 translation)

An ISP offering IPv6 connectivity has a utility of the form:

$$\Pi_6 = n_6 p_6 - D_6 n_6 \gamma_6 \,, \tag{3.3.4}$$

with n_6 the number of users choosing IPv6 connectivity at a price of p_6, and $D_6 n_6 \gamma_6$ the translation cost for those users. This expression assumes each user generates 1 unit of traffic distributed uniformly across ICPs, so that if γ_6 ICPs are not IPv6 accessible, $n_6\gamma_6$ units of traffic must be translated at a unit cost of D_6.

Public IPv4 and IPv6

An ISP offering both public IPv4 and IPv6 has a utility that is simply the sum of Eqs. (3.3.3) and (3.3.4) and is of the form:

$$\Pi_{46} = n_4 p_4 - C(n_4 - 1)_+^2 + n_6 p_6 - D_6 n_6 \gamma_6 \,. \tag{3.3.5}$$

The next subsection explains the decision mechanism of the Internet stakeholders, and the timing of those decisions.

3.3.4 Decision Mechanisms and Timing

In all scenarios, ISPs first announce a price for connectivity options, with users then choosing one in a best response fashion, *i.e.,* they select the option that maximizes their utility. ICPs decide whether or not to become IPv6 accessible in the third

and last stage of the game, again in a best response manner and based on the number of users that have chosen IPv6. ISPs are assumed aware of the rationale and economic incentives guiding users and ICPs decisions, *e.g.,* based on surveys of users and ICPs. Hence, they set prices that maximize their own profit, *i.e.,* by solving the above sequential decision process in reverse order. In the disagreement scenario, where not all ISPs agree to offer IPv6 immediately, we assume the two ISPs compete for the users by playing a best response game between themselves, and their strategies are the prices of the services. In the consensus scenario, however, the problem reduces to an internal optimization of a single ISP that offers both services, namely, IPv4 and IPv6 connectivity.

An alternate game would have users and ICPs aware of each others decisions, deciding simultaneously rather than sequentially. This assumes that users are able to predict how ICPs will respond to their decisions and vice versa, and makes for a more complex and possibly less realistic game (neither users nor ICPs may have access to the necessary information). More importantly, the outcomes are similar to those of the simpler sequential game. As a result, we focus on the latter. Another alternative scenario is when the decision making process of ICPs is in a different time scale compared to ISPs and users, *i.e.,* ICPs re-evaluate their decisions less often than ISPs and users. We relegate the analysis of this scenario to Appendix C.

3.4 Model Solution

3.4.1 Disagreement Scenario

This section considers scenario 3.2.4, which involves (two) ISPs competing for users and offering different connectivity options. One ISP relies on IPv6, but the other has deferred upgrading to IPv6. Instead, it chooses to either incur the (growing) cost of acquiring public IPv4 addresses, or to assign private IPv4 address to new users and rely on translation (CGNs) to connect them to the public Internet. Here we present the solution to the former scenario (public IPv4 vs. IPv6), and relegate the latter (private IPv4 vs. IPv6) to Appendix A.

Specifically, we assume *rational* and *myopic* ISPs that engage in a repeated multi-stage game played each time the Internet user population increases by $\delta < 1$ new users. Again, in this scenario, one ISP offers IPv6 and the other stays with public IPv4 connectivity. Public IPv4 has an edge when it comes to connectivity quality ($a_6 > 0$), but that edge is present only for the fraction γ_6 of ICPs that require translation. Conversely, the disadvantage of public IPv4 is the likely cost of acquiring additional public IPv4 addresses.

As per Eq. (3.3.1), users' utility depends on price (p_R), quality of connectivity (a_R), and the fraction γ_R of ICPs affected by quality impairment associated with connectivity option R. γ_R is assumed known to users, and in the case of IPv6 depends on the outcome of the previous round of the game, *i.e.,* how many ICPs

have become natively accessible. For tractability purposes, we assume that σ, *i.e.,* the sensitivity of users to quality impairments, is uniformly distributed in $[0, 1]$. Section 3.6 relaxes this assumption, and investigates the outcome with a more general distribution.

Hence, in round i of the game and assuming IPv6 and public IPv4 ISPs announced prices of p_6 and p_4, users and ICPs decisions proceed as follows. Based on Eq. (3.3.1), a user with quality sensitivity σ chooses IPv6 if, $1 - p_6 - \sigma a_6 \gamma_6^{(i-1)} \geq 1 - p_4$, where $\gamma_6^{(i-1)}$ denotes the fraction of ICPs not yet IPv6 accessible after round $(i-1)$ (this information is available after each round, with $\gamma_6^{(0)} = 0$ for completeness). Hence, the fraction $\sigma_6^{(i)}$ of (new and existing[30]) users choosing IPv6 in round i satisfies

$$\sigma_6^{(i)} = \begin{cases} 0 & \text{if } p_4 - p_6 < 0 \\ \frac{p_4 - p_6}{a_6 \gamma_6^{(i-1)}} & \text{if } 0 \leq p_4 - p_6 \leq a_6 \gamma_6^{(i-1)} \\ 1 & \text{if } p_4 - p_6 > a_6 \gamma_6^{(i-1)} \end{cases} . \tag{3.4.1}$$

The dependency on the price differential $p_4 - p_6$ is intuitive. For example, when IPv6 is priced higher than IPv4, IPv6 adoption is zero, while when the discount for IPv6 is larger than the quality penalty perceived by the most quality sensitive user ($\sigma = 1$), then all users select IPv6.

[30]In section 3.6 we show that our results remain qualitatively similar even if users have inertia in decision making, *e.g.,* contractual agreement, etc.

An ICP reevaluates its IPv6 adoption decision once knowing the outcome of users' decisions. Again, for tractability purposes, ICPs are assumed to derive homogenous revenues from users (*i.e.,* $\beta = 1$), but we relax this assumption in Section 3.6.

From Eq. (3.3.2), ICPs adopt IPv6 if the difference between the added revenue, $n_6 a_6$ (remember $\beta = 1$), this generates and the upgrade cost $S_{\text{infra}}\theta c_6$ is positive. The latter depends on the current size of the ICP's infrastructure, S_{infra}, which is proportional to the Internet user-base in round $(i-1)$, *i.e.,* $1+(i-1)\delta$ (where 1 is the size of the existing Internet user population). Conversely, the revenue increase created by becoming IPv6 accessible is proportional to the number of users choosing IPv6 in round i, *i.e.,* $n_6^{(i)} = (1+i\delta)\sigma_6^{(i)}$. Assuming θ is uniformly distributed in $[0,1]$ (which we again relax in Section 3.6), ICPs for which becoming IPv6 accessible yields a positive profit in round i are those with $\theta \leq \theta_6^{(i)}$ (conversely, the fraction of IPv6 accessible ICPs after round i is $\gamma_6^{(i)} = 1 - \theta_6^{(i)}$), where

$$\theta_6^{(i)} = \begin{cases} \frac{ka_6}{c_6} & \text{if } p_4 - p_6 > a_6\gamma_6^{(i-1)} \\ \frac{ka_6\sigma_6}{c_6} & \text{if } \frac{\gamma_6^{(i-1)}(1-\gamma_6^{(i-1)})c_6}{k} \leq p_4 - p_6 \leq a_6\gamma_6^{(i-1)} \\ 1-\gamma_6^{(i-1)} & \text{Otherwise} \end{cases}, \tag{3.4.2}$$

where for notation simplicity $k = \frac{1+i\delta}{1+(i-1)\delta}$ is the relative growth in user population between rounds $(i-1)$ and i.

The first expression of Eq. (3.4.2) corresponds to all users selecting IPv6, *i.e.,*

$\sigma_6^{(i)} = 1$, which yields the maximum possible adoption of IPv6 among ICPs. IPv6 adoption progressively decreases as fewer users select IPv6 (second expression), down to no less than $1 - \gamma_6^{(i-1)}$, which reflects the fact that ICPs that upgraded to IPv6 in an earlier round do not revert their decisions.

Eqs. (3.4.1) and (3.4.2) are known to the two competing ISPs, which use them to optimize their own utility functions, as expressed in Eqs. (3.3.3) and (3.3.4). This yields the following expressions for optimal prices, where for simplicity we omit the index i and use $\gamma_6 = 1 - \theta_6$.

$$p_4^* = \underset{p_4}{\text{argmax}}\{(1+i\delta)(1-\sigma_6)p_4 - C(((1+i\delta)(1-\sigma_6)) - 1)_+^2\} \tag{3.4.3}$$

$$p_6^* = \underset{p_6}{\text{argmax}}\{(1+i\delta)\sigma_6 p_6 - D_6\sigma_6(1+i\delta)\gamma_6\}\,. \tag{3.4.4}$$

The two equations are coupled through Eqs. (3.4.1) and (3.4.2).

Explicitly solving this joint optimization is difficult[31]. It can be formulated as the solution of a best response game between the ISPs, each successively announcing and reacting to the other's price. In general, the game does not have a Nash Equilibrium to which prices would converge. In particular and as illustrated in section 3.5.1, instances of "cycles" in the ISPs' search for optimal prices arise in

[31] Analytical solutions can be obtained, but are mostly negative results, *e.g.,* the absence of a Nash Equilibrium, which do not shed insight into the problem. Hence, we resort to numerical investigations to explore the solution space.

many cases. In other words, competition (or disagreement) between ISPs on the basis of connectivity makes identifying rational operating (pricing) points difficult.

Interestingly but not surprisingly, dependencies between Internet stakeholders' decisions are largely responsible for this. In particular, if ICPs' decisions were independent of those of users (or proceeded at a much slower pace), the game would typically admit a unique Nash Equilibrium (see Appendix C).

3.4.2 Consensus Scenario

A scenario where all ISPs offer IPv6 and another alternative, *e.g.,* public IPv4, is equivalent to a monopolistic ISP that serves all of the users. Choices need to be preserved, as users (and ICPs) are likely to remain heterogeneous in their willingness to accept a migration to IPv6. However, connectivity options should not be the basis on which ISPs compete. In other words, this scenario is equivalent to a pricing problem including a single provider with two types of products.

Consider an ISP offering its users (new and existing) the choice between traditional public IPv4 connectivity and IPv6 connectivity at prices of p_4 and p_6, respectively. As in the previous section, users that opt for IPv6 must undergo translation when connecting to the γ_6 fraction of ICPs that are not yet IPv6 accessible. As before, translation introduces impairments of relative magnitude a_6. Similarly, the ISP incurs a cost of D_6 per unit of traffic that needs translation. The ISP has an existing user-base of unit size, and therefore owns a unit-size block of public IPv4

addresses. If it needs additional public IPv4 addresses, it acquires them at a cost that, again as before, grows quadratically, *i.e.,* based on Eq. (3.3.3). ICPs decide to become IPv6 accessible following the same process as that of last section. We describe next how the ISP selects the prices p_4 and p_6 that maximize its profit.

Growth in the Internet user population again proceeds in steps of size δ that coincide with epochs where the ISP adjusts its prices p_4 and p_6. Choosing optimal prices involves solving the following optimization problem

$$\begin{aligned}(p_4, p_6) = \underset{(p_4,p_6)}{\operatorname{argmax}} \Bigg\{ & \\ & (1+i\delta)\left(1-\frac{p_4-p_6}{a\gamma_6^{(i-1)}}\right)p_4 - \\ & C\left(\left((1+i\delta)\left(1-\frac{p_4-p_6}{a\gamma_6^{(i-1)}}\right)\right)-1\right)_+^2 \\ & +(1+i\delta)\left(\frac{p_4-p_6}{a\gamma_6^{(i-1)}}\right)p_6 \\ & -D_6(1+i\delta)\left(\frac{p_4-p_6}{a\gamma_6^{(i-1)}}\right)(1-\theta_6^{(i)})\Bigg\}, \end{aligned} \tag{3.4.5}$$

where $\gamma_6^{(i-1)} = 1 - \theta_6^{(i-1)}$ is known, while $\theta_6^{(i)}$ needs to be replaced by its expression from Eq. (3.4.2). Note that different expressions must be used for $\theta_6^{(i)}$ depending on the value of $p_4 - p_6$. It is the need to consider those different cases that makes solving Eq. (3.4.5) cumbersome though not impossible.

Except for the fact that the optimal price for public IPv4 always satisfies $p_4 = 1$ (actually just below 1 to ensure positive utility), the expression for an explicit solution for Eq. (3.4.5) sheds little light on the role of different parameters, the reader

is referred to [60] for more details. for details, and we instead rely on numerical examples to explore the range of outcomes. The next section discusses the results of numerical analyses of the models.

3.5 Results

3.5.1 The Impact of Disagreement

ISPs' inability to converge to jointly optimal prices is primarily because the coupling between users and ICPs' decisions introduces two distinct strategies for the IPv6 ISP, and correspondingly a discontinuity in its utility function. When the price of public IPv4 connectivity is high enough, it is best for the IPv6 ISP to heavily discount its connectivity to attract many users and in turn convince many ICPs to become IPv6 accessible, which lowers translation costs. This, however, triggers a price decrease from the public IPv4 ISP to recoup part of its lost user-base, and then forces the IPv6 ISP to itself lower its price to maintain a sufficiently attractive discount. This eventually results in a public IPv4 price that is too low to allow the IPv6 ISP to give a large enough discount. The better strategy for the IPv6 ISP is then to reduce its discount and attract fewer users. Each user generates a higher revenue, and because there are few of them, translation costs are low. This pattern is shown in Fig. 3.1 that plots each ISPs' best-responses as a function of the other's price, and includes an instance of a pricing cycle.

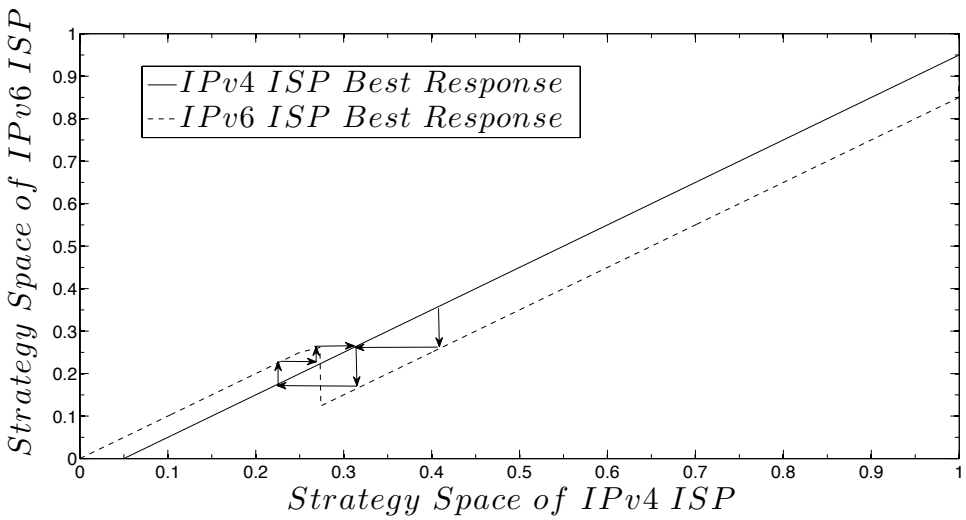


Figure 3.1: Cycle in ISPs best response game

Cycles occur when the utility gap between the two strategies of the IPv6 ISP is large enough to ensure that its best-response function and that of the IPv4 ISP do not intersect. Fig. 3.2 explores how often this arises across a reasonable range of configurations. The price of IPv4 addresses is chosen to have a normalization constant $C = 1$, so that the quadratic cost function for IPv4 addresses yields a value of 1 when the number of IPv4 Internet users reaches $n_4 = 2$, *i.e.,* doubles. In other words, doubling the size of the current IPv4 Internet yields a public IPv4 address price equal to the value of Internet connectivity itself. This choice reflects the fact that according to current statistics there were about 2 billion Internet users by the end of 2012, and given the $\approx 50 - 75\%$ utilization of the address space, a doubling of IPv4 users is then still possible. The per-user IPv6 conversion cost, c_6, is assumed to be ten percent of the base value of Internet connectivity.

Fig. 3.2 shows the outcome of the game played by the two ISPs as a function

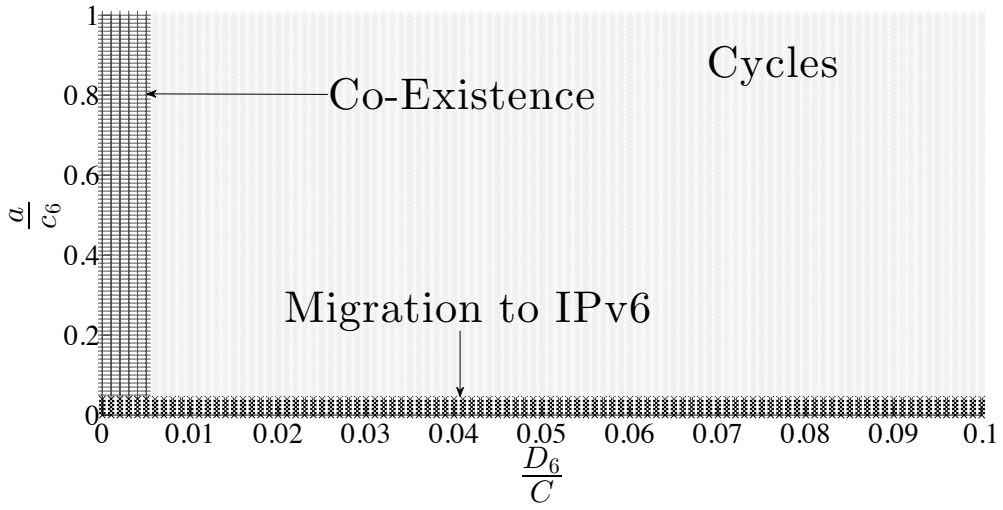


Figure 3.2: IPv6 vs. public IPv4 competition

of unit translation costs, D_6, and the non-native connectivity quality impairment a user incurs, a. D_6 is varied from zero to ten percent of the per user IPv4 address acquisition cost, while a is varied from zero to the per user IPv6 conversion cost an ICP incurs. The figure illustrates the presence of cycles in a wide range of configurations, and in particular as soon as both a and D_6 slightly grow from zero. Similar results emerge from a scenario with an IPv6 ISP and a private IPv4 ISP, but we relegate them to Appendix A. These results show that disagreement between ISPs on offering IPv6 as a connectivity option can potentially derail the current progress of IPv6 adoption, by making it hard for the Internet stakeholders to identify winning strategies.

3.5.2 The Benefit of Consensus

The previous section illustrated the difficulty of devising effective strategies, when ISPs tackle public IPv4 address shortage with competing connectivity options. The intent of this section is **not** to argue that to migrate to an IPv6 Internet, we need to forfeit competition among ISPs. This would be neither realistic nor meaningful. Instead, we want to argue for shifting competition away from connectivity choices, *i.e.,* have a consistent offering of connectivity choices among ISPs.

Unsurprisingly, IPv6 adoption and the ISP's pricing strategy are directly affected by C, the normalization constant for the cost of acquiring additional public IPv4 addresses, and D_6, the translation cost of one unit of traffic. In addition, two other parameters indirectly affect the ISP's strategy because of how they influence users and ICPs decisions, namely, c_6, the per-user cost of upgrading an ICP's infrastructure to IPv6, and a, the relative magnitude of the impairment that translation causes, and consequently the loss in quality-of-experience for users and the related revenue loss for the ICPs.

It is possible to scope the ranges some of those parameters can span, *e.g.,* $C \leq 1$, but a complete sampling of this four-dimensional space is impractical. We rely instead on several figures to report how the outcome changes as some parameters vary, while others remain fixed. The figures help identify parameters that have a significant effect on IPv6 adoption by both users and ICPs; hence suggesting possible strategies.

Figure 3.3 illustrates an intuitive outcome, and in the process offers some level of "sanity checking" of the model. In particular, it confirms the expected negative impact on ICPs' adoption of decreasing their adoption costs, c_6, while a remains constant. However, this figure alone ignores the effect of a, which as we discuss next, can have an ambivalent effect. a is the only factor that couples ISPs, ICPs and users, hence, it is important to investigate its impact on the future of IPv6 adoption.

Figs. 3.4 and 3.5 plot four quantities as a function of a, while c_6 is assumed to stay put[32]. The left hand-sides of the figures give the cumulative per-user discount the ISP offers to its users, and the ISP's total profit when the size of the Internet user population grows by 100%. The right hand-sides of the figures give, after doubling the size of the Internet user population (*i.e.,* growth by 100%), the (final) fractions of users that have opted for IPv6, and ICPs that have become IPv6 accessible. The figures report the results for two different configurations, namely, small and large values of C, and for each configuration consider different ratios between translation costs and IPv4 address acquisition costs, *i.e.,* D_6/C takes values 0.1, 1 and 10. These figures identify two parameters that have an impact on IPv6 adoption, namely, a and D_6.

Consider first the effect of a decrease in the level of impairment, a, that translation imposes. Such a decrease can (initially) make IPv6 more attractive to users

[32]Although c_6 remains constant throughout the analysis, in the above ratio it serves as a normalization factor.

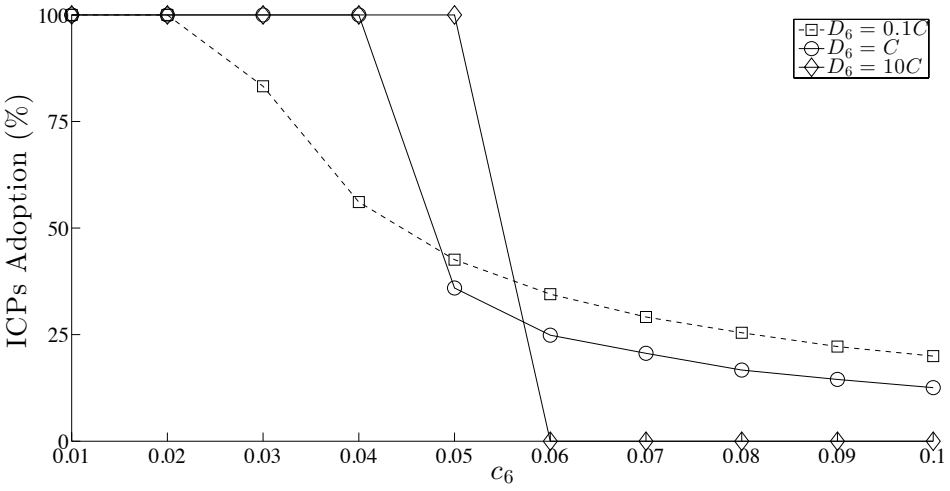


Figure 3.3: ICPs adoption levels for small C

by lowering the penalty they incur when accessing ICPs that are not yet IPv6 accessible. This can increase the number of users that choose IPv6, which can in turn entice more ICPs to become IPv6 accessible; possibly starting a positive feedback loop in IPv6 adoption. On the flip side, a lower a value also decreases the potential per-user revenue gain ICPs derive from becoming IPv6 accessible. This makes it more likely that revenue increases won't offset adoption costs; hence reducing ICPs' adoption of IPv6. This would in turn make IPv6 less attractive to users, and having fewer users opting for IPv6 would further reduce its attractiveness to ICPs. As we can see, the role of changes in a on IPv6 adoption is unclear, and the figures help elucidate under which conditions changes in a improve IPv6 adoption.

First, the figures illustrate that an increase in a (equivalently, in the ratio a/c_6) systematically results in higher IPv6 adoption by ICPs and to a lesser extent users. In the case of ICPs, a/c_6 represents the ICP's return on IPv6 adoption from a single

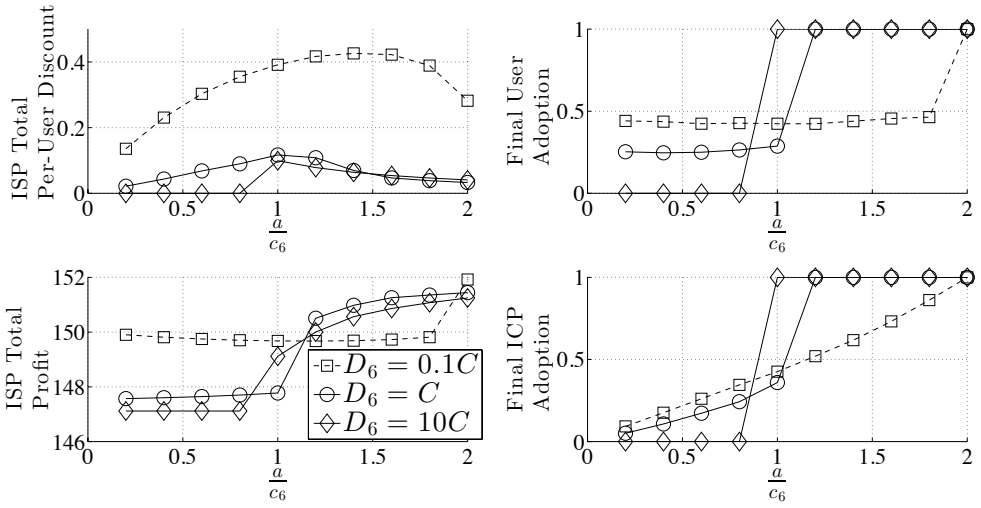


Figure 3.4: Total profit, discount & adoption levels for small C

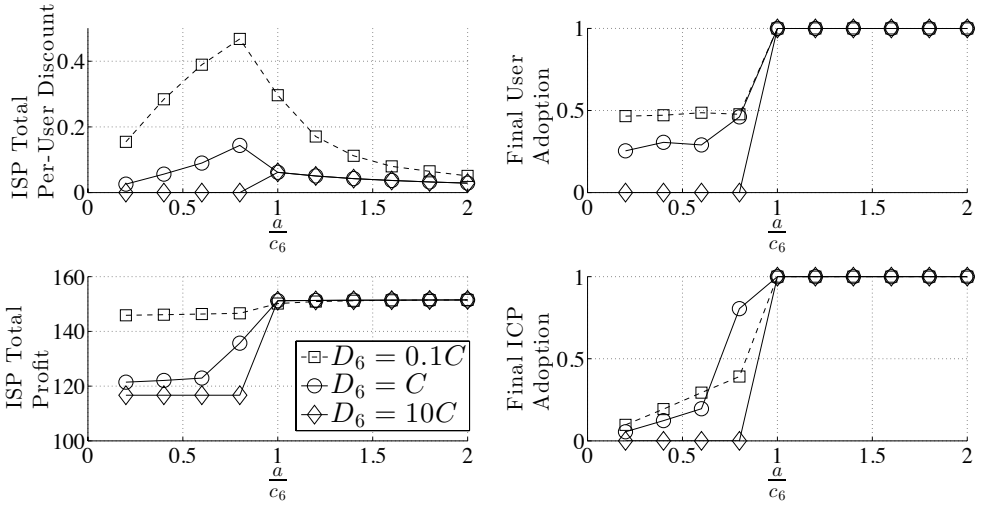


Figure 3.5: Total profit, discount & adoption levels for large C

user. An increase in this return motivates more ICPs to make such an adoption choice. When this increase is through an increase in a (rather than a decrease in c_6 that is trivially beneficial to both ICPs and users), the greater number of ICPs that opt to become IPv6 accessible offsets the larger penalty that users suffer when accessing ICPs that are not IPv6 accessible. In other words, users experience greater impairments when accessing ICPs that still require translation, but because there are fewer such ICPs, the impact is mitigated. Consequently, the number of users that choose IPv6 is not overly affected even if differences exist. In general, while both Figs. 3.4 and 3.5 establish that a larger ratio a/c_6 benefits IPv6 adoption by both ICPs and users, ensuring a complete migration to an IPv6 Internet requires a large enough value. How large this value needs to be depends on a number of factors, and in particular C and D_6, which as we discuss next introduce some interesting behaviors in their own right.

Specifically, consider scenarios in Figs. 3.4 and 3.5, for which the ICPs' return on IPv6 adoption, a/c_6, is low (less than one), *i.e.,* the ICPs have limited incentives for becoming IPv6 accessible. In such a regime, low translations costs[33], D_6, afford the ISP enough leeway to price IPv6 competitively and convince some users, and consequently ICPs, to adopt IPv6. This is reflected in the higher adoption levels of both users and ICPs as D_6 decreases. Interestingly, increasing the ICPs' return on IPv6 adoption a/c_6 has little effect on the number of users that adopt IPv6, though

[33]Note that changes in translation costs are chosen proportional to C, which explains in part the similarity of Figs. 3.4 and 3.5.

it affects (increases) the number of ICPs that elect to become IPv6 accessible. As alluded to earlier, this is because while there may be more ICPs that can be accessed natively over IPv6, this benefit is offset by the greater impairments users experience when accessing the remaining ICPs.

Further increases in the ICPs' return on IPv6 adoption (a/c_6) eventually trigger a shift in adoption, with all users and ICPs adopting IPv6. When and how this shift happens is, however, affected by the relative magnitude of IPv4 address acquisition costs, C, and IPv6 address translation costs, D_6.

When IPv4 address acquisition costs are high, and $a/c_6 > 1$ (Fig. 3.5), the shift is abrupt. This is because the high cost of IPv4 addresses entices the ISP to aggressively discount IPv6 early on to quickly convince ICPs and users alike to adopt IPv6. This can be seen by comparing the left hand-side plots of Figs. 3.4 and 3.5 that report the total discount the ISP offers to entice users to adopt IPv6, and the ISP's total profit, which shows this abrupt transition is beneficial to the ISP by saving large discounts given to users. In contrast, when IPv4 addresses costs are relatively low and $a/c_6 > 1$ (Fig. 3.4), the slow decrease in total discount (increase in total profit) can be seen to be dependent on the relative magnitude of translation costs.

In particular, with low translation costs (D_6), an ISP may initially offer only a limited discount for IPv6, which can prevent full IPv6 adoption and prolong the coexistence of IPv4-only and IPv6-only Internets (as the Internet user-base

grows, so do the benefits for ICPs of becoming IPv6 accessible, but also do their upgrade costs). In other words, if IPv4 addresses remain cheap for an extended period of time, it not only prolongs the transition to an IPv6 Internet, it may also make it significantly more expensive by deterring many ICPs from migrating early; hence, incurring higher conversion costs later on (they will need to convert a bigger infrastructure).

The next section summarizes our findings, while suggesting guidelines to avoid strategies that can derail IPv6 adoption in the future.

3.5.3 Findings

This section summarizes five of the major findings models of the previous section posit:

(i) Disagreement between ISPs on what connectivity options they offer to their users has a deteriorating effect on IPv6 adoption. As a result, ISPs should avoid *only* offering alternatives to IPv6, since it can derail the progress of IPv6 adoption.

(ii) Consensus among ISPs on offering IPv6 adoption along with other connectivity options, helps identify strategies that lead to a smooth IPv6 adoption progress. In other words, when ISPs all offer IPv6 as one of their connectivity options, devising successful migration strategies becomes easier for other major Internet stakeholders.

(iii) Decreasing the migration costs for ICPs c_6 accelerates IPv6 adoption. Tech-

nology developers should facilitate migration to IPv6 by developing more affordable IPv6 technologies.

(iv) Although the existence of translation devices (IPv4↔IPv6) is necessary for an uninterrupted migration to an IPv6 Internet, the presence of high quality translation devices (*i.e.,* small a) has a negative impact on IPv6 adoption by ICPs, since these devices facilitate a loss-less access to IPv4-only content for IPv6-only users, hence, eliminating incentives for IPv6 adoption by ICPs.

(v) Translation costs incurred by ISPs D_6 can have an ambivalent effect on IPv6 adoption: when translation impairment level (a) is small, smaller translation costs (D_6) accelerates IPv6 adoption by ICPs and users, and also contributes to higher profits for ISPs; however, when a is large, the impact of D_6 depends on the cost of IPv4 address acquisition C, when C is small, larger D_6 values are beneficial, but when C is large, D_6 has a neutral impact on IPv6 adoption.

These findings are the outcome of models from previous sections, which, for tractability and clarification, were solved with some simplifying assumptions, *e.g.,* homogenous revenue for ICPs, uniform distribution of users' sensitivity to quality, etc. In order to investigate the potential impact of these assumptions on our findings, the next section relaxes/alters them and provides the analyses for each model.

3.6 Robustness Tests

In order to capture the real world phenomena in a tractable way, every model makes some simplifying assumptions. Our modeling effort is not an exception, however, to ensure our findings are not artifacts of those assumptions, we perform a series of robustness analyses. The goal of this section is to show, through variations of the models presented in Section 3.3, that our results remain qualitatively similar to those of Section 4.5, and the findings of Section 3.5.3 hold. Next, we present a summary of these variations.

(i) Heterogeneity in revenue and cost of ICPs: In the current solution, we assume that ICPs derive revenue from users homogeneously. In reality, however, the revenues are heterogeneous, and mostly depend on the volume of an ICP's traffic, *i.e.,* popularity. Furthermore, we assume the adoption costs are uniformly distributed across ICPs, while in the real world, larger ICPs possibly benefit from economies of scale. Hence, in this variation of the model, we assume the revenue and cost of an ICP depends on its popularity.

(ii) Users' sensitivity to quality impairments (σ): In our current model, this parameter is assumed to be uniformly distributed, while in reality, the distribution is most probably not uniform. It is hard, if not impossible, to determine the exact distribution in these scenarios, however, if similar (qualitative) outcomes emerge with different distributions, the findings are on a more solid footing. Therefore, in this variation we assume users' sensitivity to quality impairments follows a different

distribution.

(iii) Users' decision making inertia: Currently, our model assumes that all users make decisions after each announcement of prices by ISPs. That might not be practical in all scenarios, *i.e.,* many ISPs impose contractual agreements on their users. Thus, in this variation of the model, we incorporate inertia in users' decisions through contractual agreements.

(iv) IPv4 address acquisition cost: In our current model, this cost is captured via a quadratic function, which reflects a scenario with growing scarcity of IPv4 addresses. However, another scenario involves ISPs with extra IPv4 addresses providing the market with enough resources to keep the cost constant. Therefore, in this variation of the model we assume the cost of IPv4 address acquisition grows linearly.

(v) The per-user cost of IPv6 adoption by ICPs (c_6): In our current model, this cost is assumed to stay put over time. However, it is more likely that it decreases as the technology matures. Hence, this variation of the model incorporates a decreasing per-user cost of IPv6 adoption.

3.6.1 Heterogeneity of ICPs

ICPs are heterogeneous in many aspects. However, it is not practical, from a modeling standpoint, to capture all of those dimensions. Moreover, from the perspective of our model, it is only the net effect of those heterogeneities that matters. Therefore,

here we only focus on one aspect of heterogeneity among ICPs, namely, popularity. We present next the changes required in the utility function of ICPs to incorporate this aspect.

The popularity is captured in Eq. (3.3.2) through parameter β, which, in previous sections, was assumed to be equal to 1 (*i.e.,* $\beta = 1$) yielding to a homogenous revenue across all ICPs. Here, we assume β is distributed in $[0, 2]$ with a certain probability distribution[34]. We assume that most ICPs are of similar popularity, but some are either highly popular or unpopular. This leads to the choice of a unimodal bell-shaped distribution function, which is adequate to show how differences in popularity of ICPs can potentially change our findings[35]. Without loss of generality, and to extend the tractability of our model, we use a unimodal Kumaraswamy distribution function for β.

Moreover, in order to incorporate the economy of scale in IPv6 adoption decisions of ICPs, we choose θ, the parameter that captures adoption cost heterogeneity across ICPs (in Eq. (3.3.2)), to be a function of their popularity. In other words, we choose θ to be a decreasing function of β, to reflect the economy of scale, *i.e.,* a more popular ICP with a larger user-base incurs (on average) lower per-user IPv6 adoption costs. Although changes in β and θ only affect the decision making process of ICPs, their incorporation in the model help capture a more general, and perhaps

[34]Choosing $[0, 2]$ instead of $[0, 1]$ only makes it easier to compare the outcome with the current outcomes, without loss of generality.

[35]Similar results emerge when we use a bimodal distribution function.

more realistic, picture of the IPv6 adoption problem. Next we show the impact of these changes on our earlier model formulation.

Formulation

Without loss of generality, we choose $\theta = \frac{2-\beta}{2+\beta}$, which bounds θ in $[0, 1]$, and is a decreasing function of β. It can be easily shown, through numerical analysis, that this choice is as good as any other decreasing function.

While the expressions for users and ISPs remain the same as Eqs. (3.4.1), (A.0.5) and (A.0.6), $\gamma_6^{(i)}$, the fraction of ICPs that do not choose IPv6 at the end of round i, changes as follows:

$$\gamma_6^{(i)} = \begin{cases} \text{if } p_4 - p_6 > a_6\gamma_6^{(i-1)} \\ \quad \mathcal{K}_{A,B}\left(\frac{\sqrt{(2+\frac{c_6}{ka_6})^2+8\frac{c_6}{ka_6}}-(2+\frac{c_6}{ka_6})}{2}\right) \\ \text{if } \frac{\gamma_6^{(i-1)}(1-\gamma_6^{(i-1)})c_6}{k} \leq p_4 - p_6 \leq a_6\gamma_6^{(i-1)} \\ \quad \mathcal{K}_{A,B}\left(\frac{\sqrt{(2+\frac{c_6}{ka_6\sigma_6})^2+8\frac{c_6}{ka_6\sigma_6}}-(2+\frac{c_6}{ka_6\sigma_6})}{2}\right) \\ \text{Otherwise} \\ \quad \gamma_6^{(i-1)} \end{cases} \tag{3.6.1}$$

where $\mathcal{K}_{A,B}(.)$ denotes the Kumaraswamy function with parameters A and B (the rest of the parameters are introduced in Eq. (3.4.2)). With the proper values for A and B, one can achieve a uni-modal distribution function in $[0, 1]$. Next, we

compare the findings from this variation of the model with our original findings.

Results

Comparing Fig. 3.6 with Fig. 3.2, one can see that this scenario yields to similar outcomes as the original scenario, *i.e.,* in the majority of the cases the disagreement between ISPs ends in a cyclical behavior. Therefore, the first finding of Section 3.5.3, *i.e.,* detrimental effect of disagreement between ISPs on technology options, holds even in the presence of heterogeneity in revenue and cost of ICPs.

The results of the consensus scenario with popularity incorporated in ICP's decisions, are labeled "Uni-modal β" in Figs. 3.7, 3.8, 3.9, and 3.10. In this scenario we assume a small C value, and $D_6 = 0.1C$, hence, we can compare them with the corresponding curve in Fig. 3.4, which is also plotted (for more convenience) in the above figures under the label "Original Scenario". A similar figure to Fig. 3.4 is relegated to Appendix B.

Comparing "Uni-modal β" curves with the "Original Scenario" (in the above figures), one can see quantitative differences, *i.e.,* the Uni-modal β scenario yields to slightly lower discounts given by ISPs, higher total profits for ISPs, and different adoption rates for users and ICPs. The main reason behind these differences is that in this scenario compared to the original scenario, more popular ICPs derive higher benefits from IPv6 adoption, therefore, their adoption rate is higher. Higher adoption rate by popular ICPs causes lower translation costs by ISPs, and therefore

lower discounts and higher profits. Users also have greater incentives to adopt IPv6 in the presence of higher adoption rate by ICPs. This is only true for small values of translation impairments (a), and for larger values of this variable most of the incentives fade away, and cause lower overall adoption rates by users and ICPs. These quantitative differences nonetheless, the outcomes are qualitatively similar, *i.e.,* the patterns are almost similar across all metrics, namely, discounts by ISPs, ISPs' profit, and adoption rates by users and ICPs. In other words, findings **(ii)** to **(v)** of Section 3.5.3 hold in the presence of heterogeneity in revenue and cost of ICPs.

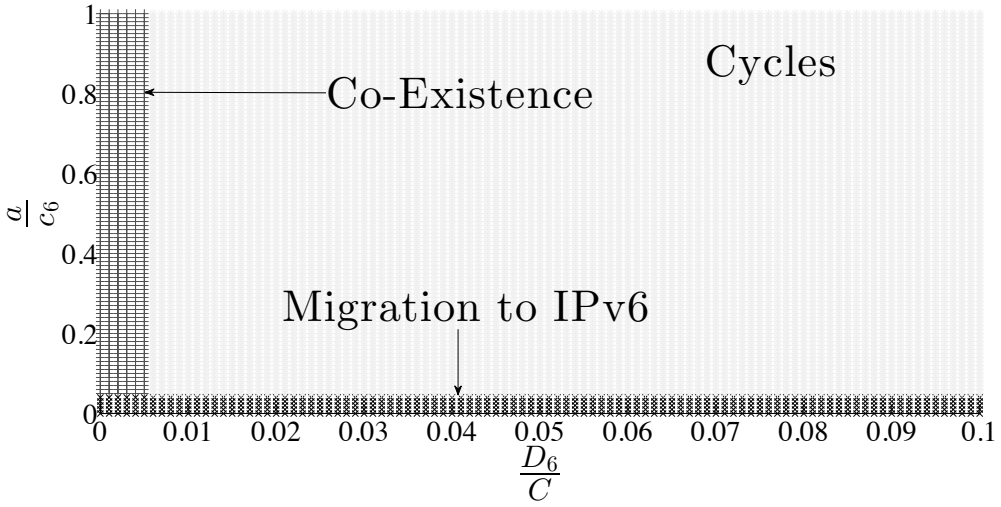


Figure 3.6: IPv6 vs. public IPv4 competition — single-modal β

3.6.2 Users Sensitivity

In the original scenario, we assumed that sensitivity to quality impairments is distributed uniformly across users. To investigate the dependency of our findings to

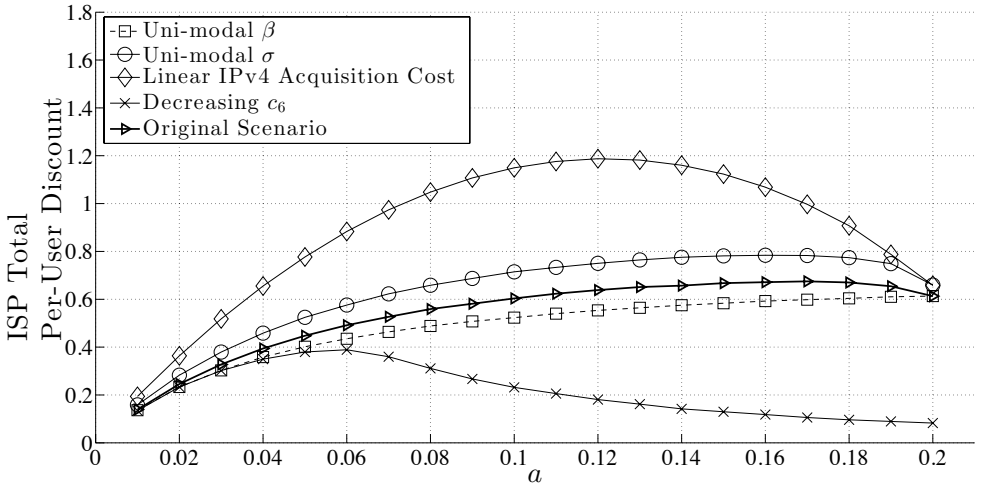


Figure 3.7: ISP's total per-user discount offered to users

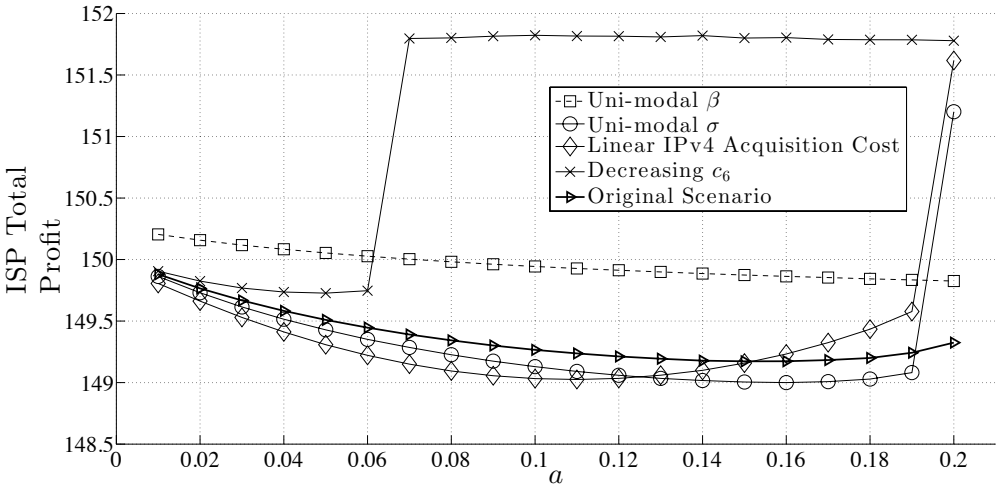


Figure 3.8: ISP's total profit

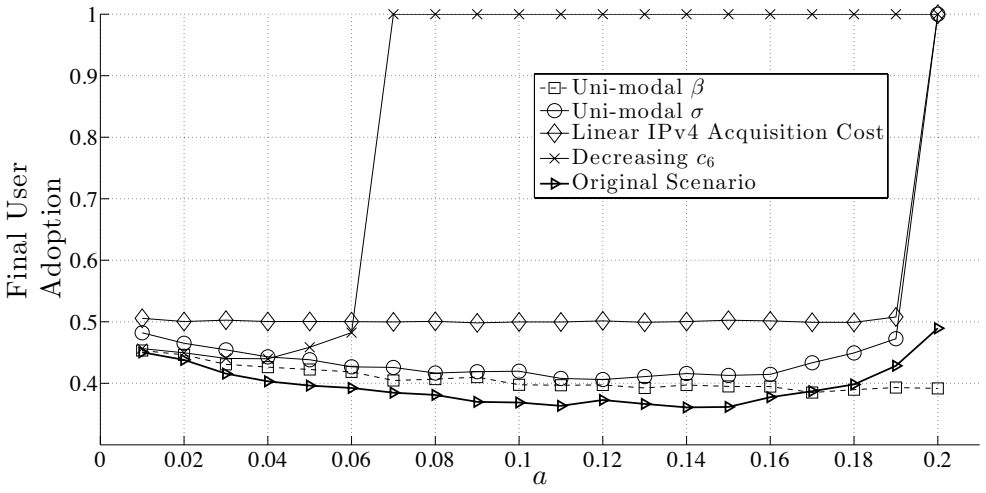


Figure 3.9: Final IPv6 adoption by users

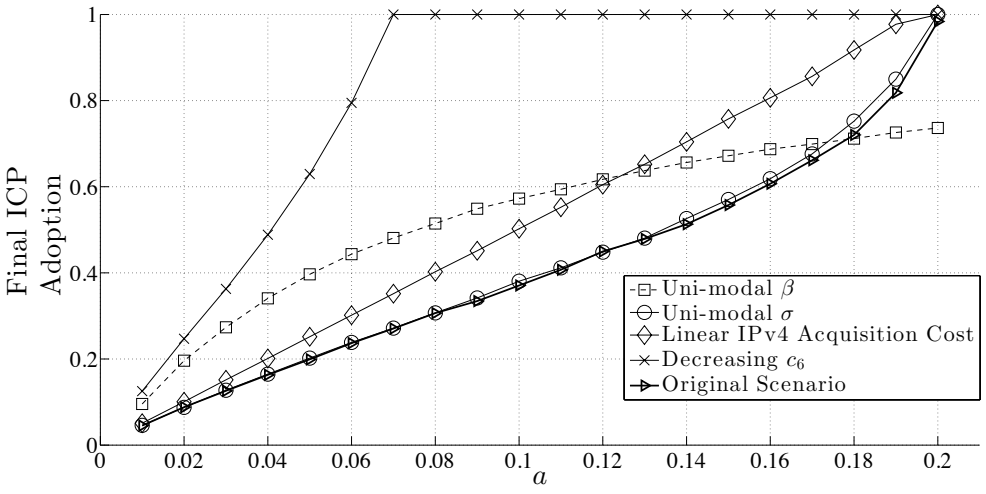


Figure 3.10: Final IPv6 adoption by ICPs

this assumption, it is best to substitute the distribution with a general distribution function. However, this seems impractical given that we rely on numerical analysis. Therefore, we only focus on a uni-modal bell-shaped distribution function, which is commonly seen in real world phenomena.

Formulation

In our model, sensitivity to quality impairments is specific to users, hence, the model formulation remains the same for ICPs and ISPs. The changes to users is reflected via σ_6 as follows:

$$\sigma_6^{(i)} = \mathcal{K}_{A,B}(\frac{p_4 - p_6}{a\gamma_6^{(i-1)}}) \tag{3.6.2}$$

which is similar to Eq. (3.4.1) except for $\mathcal{K}_{A,B}(.)$ that as alluded to earlier denotes the Kumaraswamy distribution function with parameters A and B.

Results

Similar to the previous section, comparing Fig. 3.11 and Fig. 3.2, one can see that the choice of distribution function for users' sensitivity to quality impairments (σ), does not change the outcomes qualitatively. Therefore, the first finding of Section 3.5.3 holds in the presence of a different distribution for σ.

We plot the outcomes of the consensus scenario in Figs. 3.7, 3.8, 3.9, and 3.10 under the label "Uni-modal σ". Again, in this scenario the value of C is small and $D_6 = 0.1C$ (a similar figure to Fig. 3.4 is relegated to Appendix B). Comparing

these curves with the corresponding "Original Scenario" curves, one can see slight quantitative differences, which are mainly an artifact of a different distribution for σ. However, the patterns are qualitatively similar to the original scenario. In other words, findings **(ii)** to **(v)** of Section 3.5.3 hold in the presence of a different distribution for σ.

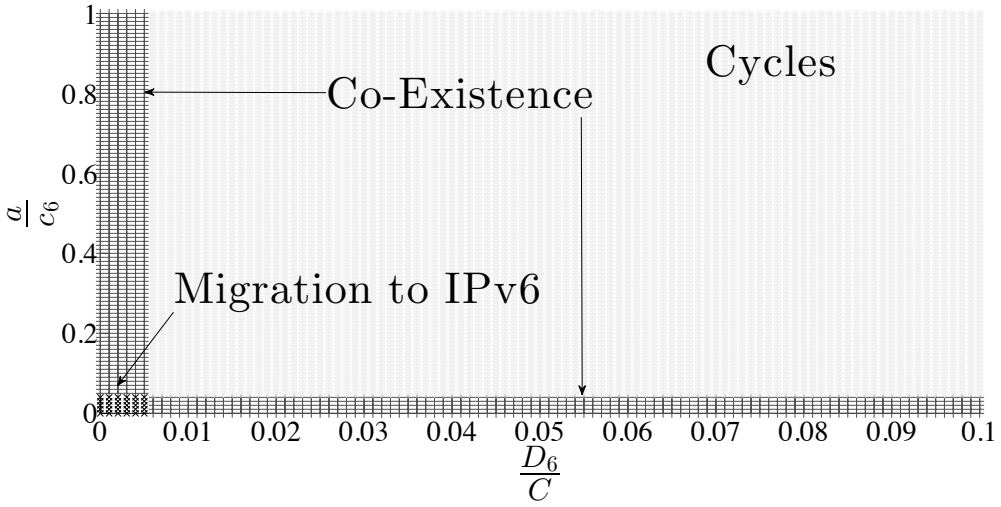


Figure 3.11: IPv6 vs. public IPv4 competition — single-modal σ

3.6.3 Users Inertia

In our initial solution we assumed all users re-evaluate their decisions after each price announcement by ISPs. In reality, this might be violated due to the inertia in decision making process of users; the most common source of this inertia is involvement in a contractual agreement, which bars users from changing their services at any time. Hence, we need to examine the validity of our findings in the presence of such inertia. In the consensus scenario, this assumption does not play a signifi-

cant role, because the monopolistic ISP can assign services to its users at any given time, in order to maximize its profit (*i.e.,* contractual agreements are internal to the ISP). However, in the scenario with disagreement among ISPs on offering IPv6, we need to analyze the model with the assumption that not all users can change their services after each announcement.

Formulation

The original formulation of the problem does not change, except that in Eqs. A.0.5, A.0.6 and 3.4.5, we substitute $(1 + i\delta)$, with a fraction of users that can make decisions at epoch i. In other words, instead of the total user population, here only a fraction of them make a decision.

Results

Fig. 3.12 demonstrates the outcome of the analysis for the above variation of the model. Comparing it with the results of our original scenario plotted in Fig. 3.2, one can see the similarity of results. In other words, this result confirms that inertia in decision making of users does not change our first finding of Section 3.5.3 regarding the disagreement scenario. The rest of the findings also remain intact, since the contractual agreements do not affect them.

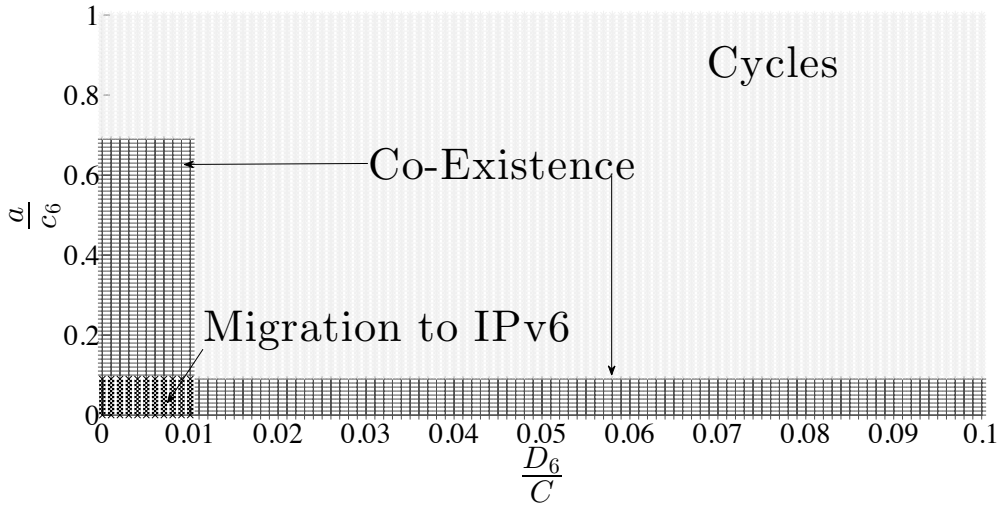


Figure 3.12: IPv6 & public IPv4 consensus — contractual agreement

3.6.4 IPv4 Address Acquisition Cost

In the original scenario, we assumed that the cost of IPv4 address acquisition grows quadratically with the number of addresses (the per address cost grows linearly). We made this assumption to capture the increasing cost of IPv4 addresses as they become scarcer. However, an alternative scenario is when ISPs with extra IPv4 addresses sell their addresses in the corresponding markets, and therefore, the supply of addresses meets the demand. In this situation, the cost of IPv4 addresses grows linearly with the number of addresses (per address cost remains constant). Here, we examine the extensibility of our findings to such scenario.

Formulation

We only need to substitute the quadratic function in the ISP's utility with a linear function. In other words, in Eqs. A.0.5 and 3.4.5, we replace $C((1+i\delta)(1-\sigma_6)-1)^2_+$

with $C((1+i\delta)(1-\sigma_6)-1)_+$.

Results

We plot the outcome of the disagreement scenario in Fig. 3.13, which is quite similar to the outcome of the original scenario in Fig. 3.2, *i.e.,* the cycles are present predominantly. Therefore, the first finding of Section 3.5.3 holds for this variation of the model.

We also plot the results of the consensus scenario in Figs. 3.7, 3.8, 3.9, and 3.10 under "Linear IPv4 Acquisition Cost". Once again, this scenario is for a small value of C and $D_6 = 0.1C$ (with a similar figure to Fig. 3.4 relegated to Appendix B). Comparing these curves with the corresponding "Original Scenario" curves, one can observe quantitative differences. However, these differences do not affect the overall patterns of adoption, *i.e.,* the patterns are qualitatively similar. Therefore, findings **(ii)** to **(v)** of Section 3.5.3 extend to this variation of the model.

3.6.5 Per-User Cost of IPv6 Adoption by ICPs

In the original scenario, we assumed that the per-user cost of IPv6 adoption for ICPs remains constant over time. However, in reality, as technology matures the cost of adoption also decreases. Here, we investigate the validity of our results in such scenario.

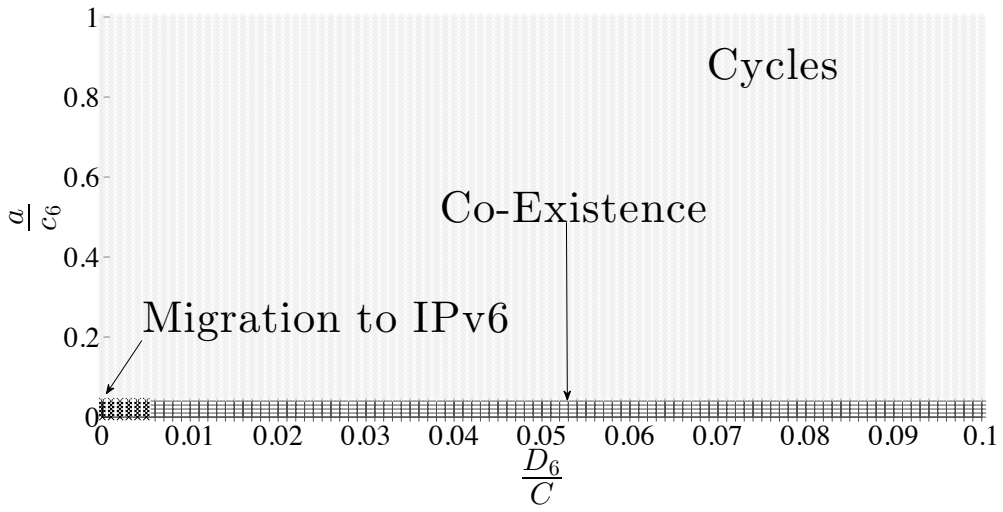


Figure 3.13: IPv6 vs. public IPv4 competition — linear IPv4 address acquisition cost

Formulation

The formulation remains the same for almost all equations, and the only difference is that we replace c_6 in Eq. (3.4.2) with $c_6(i)$, a decreasing function of i.

Results

Comparing Figs. 3.14 and Fig. 3.2 shows that temporal decrease of c_6 does not change the overall pattern of our results, *i.e.,* the first finding of Section 3.5.3.

Comparing the curves labeled "Decreasing c_6" and "Original Scenario" in Figs. 3.7, 3.8, 3.9, and 3.10, it can be easily seen that there are significant quantitative differences (a similar figure to Fig. 3.4 is relegated to Appendix B). However, this is not surprising, as decreasing the cost of IPv6 should speed up the adoption process by ICPs, which consequently causes higher profit for ISPs and higher adoption rates

for users. This exactly re-enforces what the third finding of Section 3.5.3 states. Also, the qualitative similarities between the outcomes of this scenario and the original scenario validates findings **(iv)** and **(v)** of Section 3.5.3.

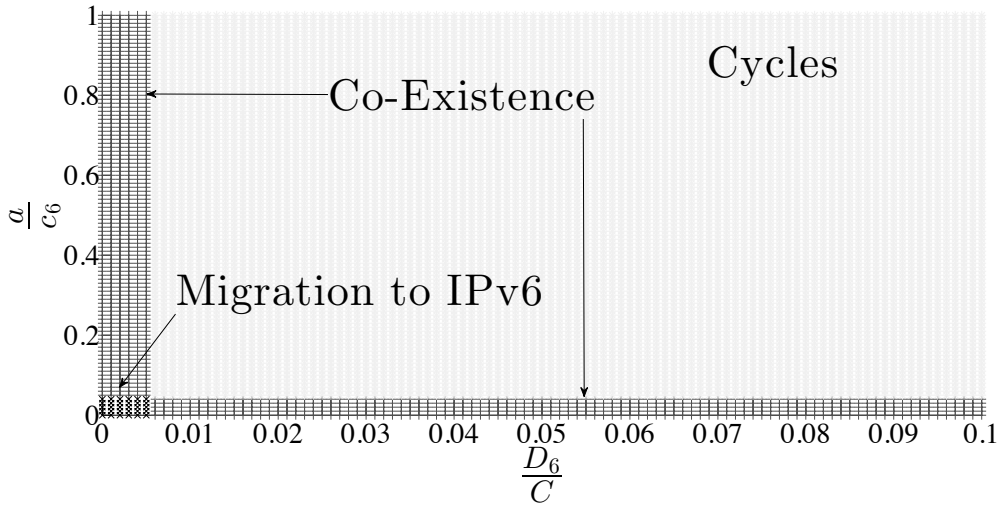


Figure 3.14: IPv6 vs. public IPv4 competition — decreasing c_6

3.7 Related Works

Explaining the slow progress of IPv6 adoption has been the focus of much prior work (see [20, 21] for a recent overview). Earlier works focused on identifying technical issues that created initial adoption hurdles [15, 57, 87, 94], but as those were eventually addressed, the attention shifted towards measuring IPv6 adoption progress [6, 8, 17, 19, 25, 40, 48, 63, 72], as well as exploring the role that economic forces may be playing [**?**, 28, 30, 34, 39, 64, 78].

Those latter works bear the most direct relevance to the investigation presented

in this chapter, with [28] echoing many of the same themes we identify, including the importance of coordination, albeit without analytical support. Casting IPv6 adoption as a game was proposed in [78], but one with Autonomous Systems as the sole players, *i.e.,* it did not account for either users or ICPs. The use of two-sided markets to capture dependencies between the decisions of Internet stake-holders was suggested in [34], but used simply to assess the impact of changing certain parameters, *i.e.,* it did not explore the possibility or competition between ISPs nor how the presence of coordinated revenue maximization by ISPs would influence the outcome.

There is also a vast literature on two-sided markets, and the reader is referred to [70, 71] for recent surveys. The most relevant works deal with competing platforms [4, 69], *e.g.,* IPv4 and IPv6, but the absence of pricing for one side of the market (the ICPs) in our context makes for a very different (and simpler) focus.

3.8 Conclusion

This chapter's work motivation is to shed light on what can potentially affect, positively or negatively, the current accelerated progress of IPv6 adoption. It proposes a number of models that capture the dependencies between different Internet stakeholders and consider various connectivity options. We therefore, explore how these dependencies and connectivity options affect the decisions of Internet stakeholders in migrating to IPv6. A first set of scenarios consider ISPs that respond to IPv4

address scarcity differently, namely, by using different connectivity options. The ISPs compete to attract users on the basis of connectivity options, and the models demonstrate how in these scenarios devising an effective IPv6 adoption strategy may be difficult, and that stakeholders can potentially derail the current progress of IPv6 adoption.

We then explore an alternative scenario in which ISPs still preserve the ability to offer multiple connectivity options, but all of them have a consensus on offering IPv6 as one of those options. While this is not by itself sufficient to maintain the current acceleration in IPv6 adoption, it affords a more predictable look at how different factors can affect the progress. In particular, it helped identify translation impairments a as an important factor that can negatively or positively affect the adoption of IPv6 in the future, *i.e.,* larger values can motivate ICPs to adopt IPv6 to avoid loss of revenue and at the same time deter users from adopting the technology, while smaller values reduces incentives for ICPs and motivates users. The model also helped identify the role IPv4 address acquisition cost C and translation cost D_6 play. In particular, it showed that low IPv4 address acquisition costs can derail the progress of IPv6 adoption at least temporarily. Even if the latter is to some extent due to the myopic decision making process of ISPs in the models, the presence of a more strategic decision does not change the outcomes significantly.

The models on which this chapter's investigation is based have numerous obvious limitations, and fail to capture the impact of many other factors. However, they

capture the interdependencies between different Internet stakeholders and their decisions. As such the results offer some insight into the potential factors that can derail or accelerate the progress of IPv6 adoption in the future.

Chapter 4

Key Factors in Protocol Adoption

4.1 Introduction

Over the past decades, the networking community has learned much about protocol design, be it in terms of performance, scalability, security, etc., or even in some cases guaranteeing the correctness of a protocol. However, we know much less about what controls a protocol's success in the "real world". IPv6 is a well-known instance, which more than two decades after its introduction still struggles to achieve wide adoption. And there are many other examples. Since 1969 the Internet Engineering Task Force (IETF) has produced over 3100 *standards track* Request for Comments (RFCs). However, in spite of a rigorous vetting process, the odds are little better than even[36] for those protocols to succeed, *i.e.,* be widely adopted by their target audience.

This raises a number of important questions that, except for RFC 5218 [76], surprisingly, have not been really addressed by previous research[37]. In particular, are specific features or properties more important than others when it comes to influencing a protocol's success? Clearly, technical correctness is important, but we have arguably made much progress in weeding out flawed protocols. External factors such as luck or commercial interests will always be present, but are unlikely to translate into systematic biases. The question is whether it is possible to carry out a quantitative and statistically rigorous investigation of protocols and protocol

[36] A random sample of close to 200 standard tracks RFCs yields a success rate of about 60%.

[37] Related works are primarily in the realm of "network economics" and therefore with a different focus than this study.

extensions[38] to identify factors with a significant influence on their success (or failure). Additionally, do these factors vary as a function of a protocol's type, *i.e.,* the functionality and environment it targets?

In this study, we apply statistical tools to mine a rich and diverse repository of protocols, namely *standards track* RFCs. Standards track RFCs correspond to protocols that have progressed through rounds of discussions in an IETF Working Group (WG), and been deemed stable and significant enough to warrant formal publication. This should, therefore, eliminate technically flawed protocols, as well as those with little community support. Our goal is to identify statistically significant features that play an important role in a protocol's success, with success defined as "broad-based" adoption among intended users. Note that in identifying such features, we do not seek to build a prediction tool. Instead, we aim to offer guidance to protocol designers by highlighting features that may be of particular significance for different types of protocols.

Our approach is three-prong. We first identify features, which reflect protocol characteristics that *may* play a role in their success. Crafting such a list is a somewhat subjective process that borrows on our experience with protocols and protocol design. We discuss our approach and its outcome in Section 4.2. Next, we construct a data set that we analyze statistically. This data set is built from a random sample of standards track RFCs (Section 4.3 discusses the selection process), which

[38]For conciseness and unless otherwise warranted, we use the term protocol to refer to both *new* protocols and extensions or new versions of *existing* protocols.

are then characterized in terms of their features as well as labeled as successful or not. Finally, as described in Section 4.4, we apply a well-established classification framework to extract protocol features that show statistically significant correlation with the success or failure of protocols. The results are then analyzed in Section 4.5 to explore their implications, and Section 4.6 performs a limited validation.The outcomes of the analysis are both intuitive and surprising. As expected, prediction accuracy remains in the $70 - 80\%$ ranges, as our focus on design features does not account for the likelihood that other non-technical factors play a role in a protocol's success. The results also confirm that markedly different features affect the success of protocols of different types. For example, while backward compatibility plays a critical role in the success of protocol extensions or new versions of existing protocols, it is of little relevance when it comes to new protocols. Other findings are, however, less immediately obvious. For example, the most significant factor contributing to the success of new application and transport layer protocols was the extent to which they were of benefit to other existing protocols. Similarly, the success of network control protocols depended heavily on their ability to realize their full value once deployed within a domain (as opposed to Internet-wide deployment). We expand on these aspects in Section 4.5.

4.2 Protocol Adoption Features

Network protocols span many functionalities, designs, and implementations. How do we capture features that set them apart, possibly influencing adoption outcomes? In this section, we put forward a nomenclature that incorporates both traditional differentiators used when describing protocols, *e.g.,* the layer they target, and aspects of value and dependencies on other protocols.

4.2.1 Characterizing Protocols

We characterize features that may play a role in a protocol's adoption along three major axes: **(i)** functionality and role; **(ii)** impact and/or dependency on other protocols; and **(iii)** primary value and how it is realized. A protocol's functionality and role affect when and where it is needed, and therefore its adoption. Similarly, how and how much a protocol interfaces to other protocols or requires them to change can make adoption harder or easier. Finally, the benefits that a protocol affords its users and when it allows them to accrue those benefits likely also plays a major role in its adoption.

In classifying protocols along those three axes, the first broadly partitions protocols according to the position (layer) they occupy in the protocol ecosystem. The second characterizes a protocol based on features that describe its interactions with other protocols, including earlier versions of itself, when applicable. Finally, the third axis reflects the functionality that motivated the protocol and its ability to

realize its value.

Before detailing the resulting list of features, we first highlight two key properties we enforce to facilitate statistical analysis. Specifically, protocol features must be:

(i) *Binary*[39]: This is to lower the "noise" inherent when measuring continuous valued functions.

(ii) *Objectively measurable*: This is again intended to limit the measurement noise that arises when subjective assessments are used to set a feature's value.

4.2.2 Features List

In investigating factors that may be influencing adoption, we consider the following twenty protocol features:

Protocol Functionality

The concept of layering has played a strong role in the design of communication protocols, and accordingly layers provide rough boundaries along which to partition protocols and their functionalities. Additional categories are, however, necessary to distinguish between protocols with similar functionality but targeting different users, *e.g.,* end-users vs. the network itself. This led to the formulation of the following six features or categories under which to classify protocols:

(1) Application (A): Protocols that provide different means of network communication to users, *e.g.,* ssh;

[39]Categorical features (*e.g.,* protocol type) can be transformed to binary features.

(2) Transport (T): Protocols that deal with end-to-end connectivity functionality, *e.g.,* TCP;

(3) Network Services (S): Protocols that target services that facilitate network use, *e.g.,* DNS;

(4) Network Control Plane (C): Protocols that affect network configuration, *e.g.,* RSVP;

(5) Network Routing Plane (R): Protocols that determine packet forwarding, *e.g.,* BGP;

(6) Network Data Plane (D): Protocols that deal with packet format and handling, *e.g.,* IPSEC.

Note that this classification omits Link and Physical layer protocols. The primary reason is that the IETF targets very few protocols in these categories.

Impact or Dependency on Other Protocols

These features capture the impact or dependency of a protocol on other protocols or network components, as well as its interaction with them. The information needed to identify them is available from the protocol's RFC and accompanying documents.

(7) New protocol vs. extension/version of an existing protocol: A new protocol is a clean design, while an extension/version of a protocol inherits and/or builds on its predecessor. This creates different adoption challenges, *e.g.,* extensions/versions need to interact with an installed-base, while new protocols

may need to displace a functionally similar protocol.

(8) Replacing another protocol: This applies to new protocols seeking to replace an existing protocol. Migrating the incumbent's user-base can be a challenge.

(9) Requiring changes to other protocols: At times, a protocol requires changes to other protocols, *e.g.,* TCP's Explicit Congestion Notification calls for changes in the IP header. Effecting those changes creates additional adoption hurdles.

(10) Generating additional value for other protocols: In some cases, a protocol adds value to other protocols, *e.g.,* IPSEC offers security to upper layer protocols. This can facilitate adoption.

(11) Affecting network (hardware or software) components: Support for a protocol can occasionally require changes to network components, *e.g.,* IPv6 affects router hardware and software. Deploying those changes can delay adoption.

(12) Backward compatibility: This applies to protocol extensions/versions and reflects whether it can interoperate with earlier versions, *e.g.,* TCP SACK. The extent to which this holds will likely affect adoption.

Value and its realization

The type of value or functionality behind a protocol's standardization can also be expected to play a role in its adoption. We distinguish between three main cate-

gories, namely, performance, security, and scalability, plus one "catch-all" category.

(13) Performance: This covers features that seek to improve communication throughput or latency, *e.g.,* as some TCP extensions arguably have.

(14) Security: This includes authentication and encryption aspects, as well as protocol mechanisms aimed at strengthening communication integrity.

(15) Scalability: Dealing with the growth of the Internet is a key feature in many protocols, which this seeks to capture.

(16) Others: As alluded to, this is a place-holder for motivations that do not belong to one of the above three categories.

While the above features identify where the value of a protocol might lie, another important aspect is realizing this value. In particular, value realization may depend on the level of adoption of a protocol. For example, does value grow with adoption, and if yes, how? We distinguish between four different scenarios.

(17) Local: Full value is realized even under limited (individual) adoption. Mobile IP can, for example, be argued to fall in this category. Such a feature is expected to facilitate adoption.

(18) Domain-wide: Adoption within the realm of a single management entity, *e.g.,* an Autonomous System, is sufficient to unlock the protocol's value. This is common with intra-domain routing protocols.

(19) Internet-wide: Realizing the bulk of the protocol's value calls for widespread adoption, *e.g.,* as is the case with IPv6. Such a constraint can be expected to make

adoption more challenging.

(20) Increasing: The value of many protocols increases with adoption, *e.g.,* the benefits of DNSSEC grow, the more widely adopted it is. The ability to progressively realize value may foster adoption.

4.3 Data Collection

The IETF relies on various vehicles to discuss and distribute protocols it seeks to standardize. Proposed IETF protocols that have reached a certain level of maturity are typically disseminated through Requests for Comments (RFCs), namely, have progressed through rounds of discussions in an IETF Working Group (WG), and been deemed stable and significant enough to warrant formal publication. RFCs can belong to different categories, and we focus on standards track RFCs. Our motivation is that they correspond to protocols that are of sound design, so that an eventual failure to gain widespread adoption is unlikely to stem from fundamental technical flaws.

We select a representative sample of RFCs in three steps. The first step involves eliminating all RFCs issued since 2009. This is to ensure that enough time had elapsed since the protocol's initial release for a reasonable assessment of its adoption status. The second step involves eliminating irrelevant RFCs from the above set. Irrelevant RFCs are those that are not introducing a new protocol or an extension to a protocol, or are associated with data link and physical layers. The reason we do

not consider data link or physical layer protocols is that IETF was not involved in the standardization of many of the protocols in these two categories, and therefore, it is unlikely that we can infer anything from the available data. The set of relevant pre-2009 standards track RFCs is still huge, and characterizing all of them (using the characteristics defined in the previous section) is infeasible, therefore, we need to find a representative sample to infer (potential) statistical correlation between a number of characteristics and the success of protocols. We seek not only to find such correlations, but to quantify them and find their statistical significance. Hence, our sample needs to be representative of all types of protocols, regardless of their popularity, to avoid potential bias in our statistical findings. In other words, our focus is on finding a sample that is not dominated by popular protocols and the rules that govern them, so that our findings can be generalized to all kinds of protocols. Thus, the third step focuses on producing a subset of RFCs for our analysis. This was performed by combining two different sampling methods, one producing 110 and the other 120 RFCs.

The motivation for using two different sampling methods is to avoid over-sampling "popular" protocols that tend to see many extensions and correspondingly generate a large number of RFCs, *i.e.,* a simple random sampling tends to sample popular protocols more often than other protocols, thereby introducing a bias in our dataset. The first sampling method involves randomly sampling all past and present IETF WGs and retaining one major pre-2009 RFC from each of the

sampled WGs (in the rest of the chapter, we refer to this sample as "WG-based"), producing 120 "major" RFCs. In other words, we select 120 WGs from the total of 179 WGs that have at least one relevant pre-2009 standards track RFC. Then, from each WG we select the major RFC it produced, *i.e.,* an RFC that is the most important one among all the standards track RFCs of that particular WG. This sampling method gives each major RFC a chance of $\frac{120}{179} \simeq 67\%$ to be included in our final dataset. The second sampling method involves selecting 110 RFCs randomly from the non-major RFCs, *i.e.,* any other relevant pre-2009 standards track RFC that is not in the 179 major RFCs set, and in the rest of this chapter we refer to the resulting sample as "random". Since there are a total of 1473 relevant RFCs that are in the standards track and pre-2009, the second sampling method basically selects 110 out of 1473-179, giving each RFC in this set a chance of $\frac{110}{1473-179} \simeq 8.5\%$.

We seek to combine these two sample sets to create a larger set that includes all types of protocols regardless of their popularities. However, we can only combine these datasets if their statistical differences are not beyond what we intended. In particular, our goal is to create a sample that is not biased towards popular protocols, *i.e.,* not containing many extensions of a few popular protocols. Therefore, we used the WG-based method to sample more "new" protocols, *i.e.,* protocols with a fresh design, even if they are not popular. Except for this feature that we intended to have, and the ones that are closely associated with it, we want the two samples (*i.e.,* random and WG-based) to have similar distributions, so that combining them

into a larger dataset does not create a new statistical bias. In other words, we want to combine the two datasets if and only if their distributions are not drastically different (except for those dimensions that we intended to be different, *e.g.,* characteristics related to new protocols). We relegate the tests that shows this similarity to the end of this section, and here only explain what are the general characteristics of the final sample.

The final sample is found by mixing the above two samples, and as a result of our two sampling methods, each relevant RFC has a non-zero (*i.e.,* more than 8.5%) chance of being a part of this mixture. If we have found all 230 RFCs in our list by simple random sampling, the chance of each relevant RFC being selected would be exactly $\frac{230}{1473} = 16\%$. However, our sampling assigns a higher chance (67%) to major RFCs, and a lower chance (8.5%) to others. As alluded to earlier, this is done to avoid the bias (towards more popular protocols) associated with simple random sampling. In other words, our final dataset is a more representative sample (compared to a simple random sample of the same size) of all types of RFCs regardless of their popularity, since it captures a larger range of protocols, and therefore, we can generalize our findings to a broader set of RFCs.

The next and most time consuming step involved characterizing the protocol described in each RFC according to the features of Section 4.2, as well as label it as successful or not. Both are essentially manual processes that involve our own experience with protocols, together with extensive reliance on a broad range of

external sources, *e.g.,* books, technology blogs, source code forums, product web pages, IETF mailing lists, etc. This process, while lengthy, was to some extent simplified by the fact that we dealt primarily with binary decisions for each feature. This mitigated the impact of unavoidable inaccuracy in the information that led to classifying each protocol as either having or not having a particular feature. Internal cross-validation was performed between the authors, and the result of our classification is available for external review (See Appendix D). The results are in the form of a spreadsheet with each row representing one of the 230 protocols under consideration, and the columns corresponding to the different features of Section 4.2 in addition to the label of successful vs. not successful. Comments are available for cells in the spreadsheet that highlight the motivations behind its setting and/or pointers to documentation used to support the choice that was made.

As protocol functionality arguably represents a natural partitioning, we present in Table 4.1 classification results for functionality features, as well as the corresponding number of protocols deemed successful in our two samples, random and WG-based. Complete results are available from the spreadsheet.

	A	T	S	C	R	D	Succ.
Random (110)	33	10	26	8	24	9	71
WG-based (120)	32	6	36	15	8	23	64

Table 4.1: Protocol classification statistics

As alluded to earlier, it is important to examine the similarity of our two datasets

before combining them into a larger dataset. We examined the similarity between the distributions of our two samples, namely, WG-based and random, by applying the binomial proportion exact test (*i.e.,* Fisher's test), each feature. In other words, for each feature we have two vectors of binary values (with sizes of 110 and 120), to which we apply the Fisher's test. The null hypothesis of this test is whether the two samples are drawn from similar distributions ($H_0 : p_1 = p_2$), since it is easier to define such a test compared to a null hypothesis that the two distributions are different. If the outcomes show that we cannot reject the null hypothesis for all of the features (except for those that we intended to be different), then it is reasonable to assume that combining the two datasets does not create a new bias. In other words, the ideal outcome is that we cannot reject the null hypothesis for most of the features with a significance level of 95%. The result of the Fisher's test is a $p-value$ that basically shows the probability of the two samples being drawn from the same distribution, *i.e.,* the significance with which the null hypothesis can be rejected is found by "1 - *p-value*". These $p-values$ are listed for each feature in Table 4.2. The rows labeled "Feature" list the corresponding feature on which the test was performed, and the rows below them (labeled "p-value") show the p-values of the Fisher's test.

It can be seen that as of the RFCs' labels, *i.e.,* being successful or not, we cannot reject the null hypothesis with 95% significance ($p-value = 0.107$). In other words, we cannot claim that the two sampling methods produce datasets that have

different distributions when it comes to success or failure of protocols. According to the results listed in Table 4.2 the above claim is true for all other features except for 4 of them, namely, being a new protocol, generating additional value for other protocols, being a Routing protocol, and being a Data plane protocol.

Among these 4 features, "being a new protocol" is the most important, since the rest of the differences are mostly associated with it (see below). By examining the datasets, we realized that the WG-based sampling produces more new protocols than the random sampling method, which as alluded to earlier, is what we expected to happen, *i.e.,* more protocols with fresh designs are going to be present in this dataset compared to the random sample.

The rest of the differences are mainly associated with being a new protocol and how the RFCs are distributed. For instance, new protocols mainly generate additional value for other protocols, compared to extensions of existing protocols that mostly generate value for their parent protocols. Therefore, since the WG-based sampling generates more new protocols, the "generating additional value for other protocols" feature has a higher proportion in this dataset compared to the random sample. Also, the difference between the proportion of the routing protocols in the two samples comes from the fact that there are a few new routing protocols in the IETF standards, and they are mainly popular, therefore, the random sampling selects more of them (*i.e.,* extensions of those popular protocols) compared to the WG-based sampling. Finally, the difference between the proportion of the Data

plane protocols between the two samples is due to having a large number of major data plane RFCs (fresh designs) from various WGs, which results in more of them being selected by the WG-based sampling.

The results of the above tests show that despite the different sampling methods, and the slight differences in the datasets that they produce (which were intentional to avoid the bias toward more popular protocols), the two datasets, namely, WG-based and random, do not have drastically different distributions. Moreover, as we explain in the next section, we categorize RFCs based on their functionalities (and being new or not for some categories) before applying the classification methods to them, and as a result, the slight differences mentioned above do not affect our findings. In other words, not only the mixture is a more representative sample of all types of protocols (regardless of their popularities), but also, according to the results of the Fisher's test, it does not introduce a new bias that affects our final findings.

Feature	Succ.	A	T	S	C	R	D
p-value	0.107	0.660	0.301	0.301	0.271	0.001	0.021
Feature	New	Replace	backwards	Net. Gear	Local	Domain	Internet
p-value	0.000	0.107	0.186	0.891	0.445	0.884	1.000
Feature	Increase	Change	Gen Value	Secur.	Scala.	Perf.	Other
p-value	0.787	0.787	0.013	0.623	0.350	0.508	0.515

Table 4.2: Results of Binomial Proportion Tests for Sample Distribution Similarity

4.4 Methodology

Given our relatively small dataset and our goal to identify features that play a major role in a protocol's adoption, we considered statistical methods such as logistic regression, decision trees, or Logistic Model Trees (LMT), instead of less-transparent and more data demanding algorithms such as neural networks or K-nearest neighbor [89]. Among those initial choices, we finally settled on binary logistic regression[40], since our focus is on the interpretability of the results (*i.e.,* regression is more transparent in terms of which factors/features are significant, and in quantifying their relative effects).

The relatively large number of features on which we rely to characterize protocols (see Section 4.2), leads us to first use stepwise regression to isolate features (a model) with the highest classification impact. Stepwise regression adds features one-by-one in its forward mode and removes them, also one-by-one, in its backward mode, and at each step examines whether a particular criterion, *e.g.,* Akaike Information Criterion (AIC), or Bayesian Information Criterion (BIC) [10], is minimized[41]. In our analysis, we focus on AIC (essentially a measure of the model's quality based on a trade-off between its goodness of fit and its complexity), as it proved the most

[40] "The binary logistic model is used to predict a binary response based on one or more predictor variables (features), making it a probabilistic classification model in the parlance of machine learning" [1].

[41] We also considered p-value thresholds, which enters (removes) a feature to (from) the model only if its significance meets the "enter" or "leave" thresholds.

efficient in selecting a model. Results for BIC and other criteria were, however, qualitatively similar.

Once key features have been identified, we feed them to the binary logistic regression method. In our investigation, a positive (negative) outcome of this method corresponds to an RFC classified as successful (unsuccessful). We rely on JMP 12 (`http://www.jmp.com`) for stepwise regression and binary logistic regression, and for the statistics it provides to characterize the outcome of the classification. Weka 3-7-12 (`http://www.cs.waikato.ac.nz/ml/weka/`) is used in turn for cross-validation (see below) as well as to generate a confusion matrix. Because of our relatively small data set (and/or large number of features), we can face a quasi-separation problem. In other words, the model overfits to the dataset, meaning that it memorizes the data instead of learning the relationship between the response and the features, and therefore, the model coefficients are mostly not statistically significant, which makes the model provide less value in terms of classification. When faced with such a scenario, we rely on built-in regularization tools provided by Weka and JMP, namely, Ridge Regression and Firth Bias-Adjusted Regression, to avoid overfitting.

The outcomes identify which features play important roles in a protocol's success or failure through two main metrics, odds ratio and statistical significance (*i.e.,* p-value). The odds ratio captures the odds that an outcome occurs given the presence of a particular feature, compared to the odds of the outcome occurring

in the absence of that feature — odds ratio values less (greater) than 1 imply the existence of a negative (positive) correlation. The statistical significance of each feature is characterized through a likelihood ratio test. This test compares the model's likelihood with that of an alternative model from which the feature is absent. The p-value is then obtained assuming a χ^2 distribution for the test statistic. The smaller the p-value, the less likely the alternative model from which the feature is absent[42].

The likelihood ratio test is based on a number of assumptions, namely, samples must be independent and identically distributed (i.i.d.), and the sample size must be large. The first two conditions (i.i.d) are met in our data, however, the last one is not, *i.e.,* we are dealing with small data. Having a large sample is required to make the approximation of the test statistics follow a χ^2 distribution, according to Wilk's theorem [88], and only if those assumptions are met the resulting p-values are accurate. The theorem shows that using the Central Limit Theorem (CLT) we can ensure the normality of our test statistic, and then using an approximation with an error of $O(\frac{1}{\sqrt{n}})$ (where n is the sample size), we can find the p-value using the χ^2 distribution [7]. Therefore, the measure of large sample here is whether the error of $O(\frac{1}{\sqrt{n}})$ is small enough to be ignored in the chi-squared distribution approximation, *i.e.,* whether the distribution is asymptotically χ^2. Since our data is being categorized later to 7 categories (see section 4.5), and each category has

[42]We target a significance $\geq 95\%$, *i.e.,* a p-value ≤ 0.05.

about 30 samples, therefore, the error term is multiplied by a factor of $\frac{1}{\sqrt{30}} \simeq \frac{1}{5}$. Even though the exact error also depends on other terms in $O(\frac{1}{\sqrt{n}})$ (where n is the size of the sample fed to the logistic regression model), the approximation is potentially not accurate. However, there are studies [7,52,74,92] that show not only there are ways of mitigating this error, but also the likelihood ratio test is relatively accurate in certain situations. Among these studies, [66, 81, 84] suggest with 5 to 20 samples per predictor variable, the asymptotic approximations become close to accurate, *i.e.,* the likelihood ratio test gives close-to-accurate p-values.

The settings in these studies are applicable to our case, since the only assumptions they have are as follows: (i) the outcomes are binary (therefore the underlying distribution under the null hypothesis is binomial or multinomial), (ii) the sample size is small (between 5 to 40), and (iii) the samples are i.i.d. Therefore, we can reasonably assume that, in our study, applying logistic regression to the datasets and obtaining the p-values using the chi-squared approximation does not yield to completely wrong or far-from-accurate outcomes.

However, we also use another approach called Exact Logistic Regression (ELR) [36,55], which basically gives exact p-values for small datasets. This method is based on permutational distributions of sufficient statistics, and does not have any asymptotic conditions or requirements The only assumption of this method is (except for i.i.d. samples) that we are interested in the inferences of certain parameters and consider the "intercept" as a nuisance parameter. Then, ELR finds the conditional

likelihood function of the "parameters of interest", and using a permutational distribution of their sufficient statistics, calculates the exact inferences (*i.e.,* p-values). This method was not used until recently due to the lack of efficient algorithms to compute the exact p-values, however, with the development of better algorithms, it is now considered a popular method (even though many statisticians regard it as too conservative in terms of finding p-values). Even though this method is an exact way of obtaining the p-values, we still report the results of the unconditional logistic regression method, since it is almost infeasible (using the machines and algorithms that are available), to apply the ELR method to datasets with (roughly) more than 60 samples and large number of features (it requires more than 2^{60} operations).

The odds ratio and p-value reflect our focus on identifying features likely to play an important role in a protocol's success, rather than develop a "predictor" for a protocol's eventual success. However, we also consider metrics that evaluate the model's predictive accuracy.

The first is the 5-fold cross-validation accuracy [50]. 5-fold cross-validation randomly divides the dataset into 5 equal subsets, and uses each subset as a test set, with the other four used to train the classifier. The classification rates on each test set are averaged to build the 5-fold cross-validation classification accuracy.

Another relevant metric is the "confusion matrix" that consists of True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP) rates. It measures the classifier's ability to correctly identify successful and unsuccessful pro-

tocols. TP (TN) is the fraction of successful (unsuccessful) RFCs properly classified. Conversely, FP (FN) is the fraction of unsuccessful (successful) RFCs classified as successful (unsuccessful).

4.5 Results

This section reports results from applying the classification of Section 4.4 to the 230 RFCs in our data set.

The (random) sampling process that generated our 230 RFCs, arguably resulted in a disparate set of protocols. This reflects protocol diversity, but makes it unlikely that the same features are behind the success (or failure) of each one of those 230 protocols. This begs the question of whether seeking to identify a common set of features is meaningful in the first place. And if not, how should we instead group protocols?

We explored this issue by first applying our classifier to the full set of 230 protocols, and then separately to new and extensions or new versions of existing protocols. The results are presented in Tables 4.3, 4.4, and 4.5, respectively, with rows listing the features identified by the stepwise regression, and for each feature highlighting its odds ratio (OR) and statistical significance (p-value) obtained through unconditional logistic regression. The reason we do not apply ELR to these datasets is their large sizes and the large number of features. For instance, for the case of all RFCs, applying ELR requires about 2^{230} operations, which is simply not feasible.

The last two rows report the 5-fold cross validation accuracy (Accuracy) and the confusion matrix. Stepwise regression generated a relatively large set (10) of "relevant" features when applied to all protocols, but only 5 features for existing and new protocols. The large number of features for these three categories is consistent with our expectations given the underlying protocol diversity, and makes interpreting the results difficult.

Odds ratios were middling (mostly in the 2-10 range for positive correlations, and similarly for negative correlations), but improved slightly when separating protocols into new and existing versions. A similar pattern was observed for p-values. A few values fell below the target 95% confidence, and separating protocols into new and existing again produced minor improvements.

This motivated grouping protocols into more consistent sets, whose success would then more likely depend on similar features. A natural grouping is along a protocol's functionality, or closely aligned with it[43]. We rely on such a grouping, and report next the results of our classification for each group. Results for the first two groups are split between new and existing protocols, but aggregated for the last three groups primarily because they include too few protocols (something that should be addressed in future works). The results of the unconditional (or Firth Bias-Adjusted[44]) logistic regression and the exact logistic regression (ELR)

[43]As a sanity check, "arbitrary" groupings were investigated, and consistently yielded poorer outcomes.

[44]As alluded to earlier, when facing complete or quasi-complete separation problem in our data,

Feature	OR	p-value
Protocol type=A	2.17	0.054
Protocol type=R	3.11	0.038
Backward compatible	2.75	0.001
Affect net. gear	0.20	0.001
DomainWide	2.28	0.060
Changes Other Protocols	0.21	0.020
Gen. value for other protocols	3.47	0.000
Security motivated	0.50	0.074
Scalability	3.21	0.064
Performance motivated	6.28	0.002
Accuracy	67.0%	
Confusion Matrix	TP=0.79 FN=0.21 TN=0.50 FP= 0.50	

Table 4.3: All protocols (230 RFCs)

are presented in the form of tables with rows listing the features identified by stepwise regression, and for each feature highlighting its odds ratio (OR_u, obtained from the unconditional or Firth Bias-Adjusted logistic regression method), approximated statistical significance ("L-R p-value", obtained from the Likelihood Ratio

we resort to Firth Bias-Adjusted or Ridge Regression method instead of the unconditional logistic regression method.

Feature	OR	p-value
Protocol type=A	5.19	0.019
DomainWide	4.27	0.059
Changes Other Protocols	0.09	0.032
Gen. value for other protocols	19.62	0.000
Performance	9.51	0.060
Accuracy	76.9%	
Confusion Matrix	TP= 0.72 FN= 0.28 TN=0.82 FP= 0.18	

Table 4.4: New protocols (78 RFCs)

Feature	OR	p-value
Protocol type=R	6.83	0.002
Backward compatible	3.65	0.011
Affect net. gear	0.13	0.000
Security	0.28	0.010
Performance	9.46	0.006
Accuracy	71.7%	
Confusion Matrix	TP= 0.85 FN= 0.15 TN=0.48 FP= 0.52	

Table 4.5: Existing protocols (152 RFCs)

test), conditional odds ratio (OR_c, obtained from the ELR method), and the exact statistical significance ("Exact p-value", obtained from the ELR method). The last two rows report the 5-fold cross validation accuracy (Accuracy) and the confusion matrix.

Finally, we also perform another test using the ELR method to roughly compare its outcomes with the stepwise regression method (even if this is not a fair comparison). The test involves using each of the features in a single-variable model and investigate its correlation with success or failure of the protocols. This is basically examining the impact of each feature on success of a protocol, regardless of other features. The outcomes are $p-values$ that show the significance of those potential correlations, and if the features that are found to be significantly correlated match those identified by the stepwise regression, we have a rough validation of the stepwise regression's outcome. We also report on the results of this test for each of our categories.

4.5.1 Application & Transport Layer Protocols

Our first group combines application and transport layer protocols. They share many properties, *e.g.,* both reside primarily in end-systems, but more particularly, among the twenty features we consider, they have in common the fact that their value usually increases with adoption. This is rarely the case with other types of protocols. Results are reported in Table 4.6 for new protocols and Table 4.7 for

existing protocols.

	Uncond. Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Gen. value for others	24.00	0.001	20.97	0.004
Accuracy	78.6%			
Confusion Matrix	TP=0.67 FN=0.33 TN=0.92 FP= 0.08			

Table 4.6: New application & transport protocols (28 RFCs)

Table 4.6 highlights a point that adding value to other protocols is the single most important factor behind the success of a new application or transport layer protocol. In hindsight, this may be intuitive for transport protocols which need to demonstrate value to applications (and application protocols) to be adopted. This is less so for application layer protocols, though many still end-up interacting with other application layer protocols; in the process potentially contributing value to those protocols. Note also that the importance of this feature does not mean that it is a necessary condition for a protocol's success. As a matter of fact, the high FN value indicates that close to 30% of successful protocols did not have this property. Nevertheless, the results indicate that this feature plays an important role in the eventual success of a new application or transport protocol.

It can be seen that the odds ratios obtained from the unconditional logistic regression and the exact logistic regression are close, and the approximation of the

p-values obtained from the likelihood ratio test is also close to the exact p-values. This highlights the fact that our initial approach of approximating the p-values does not create a large error, even though we are dealing with small data, which is aligned with the findings of the other studies (note that we have about 28 events per predictor variable). Another finding is that the $p - values$ generated by the ELR method also only identify "Generating value for other protocols" as significant. This is aligned with the outcome of the stepwise regression method we use to select features.

	Uncond. Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Backwards compatibility	15.59	0.003	13.66	0.012
Security	0.17	0.038	0.18	0.096
Accuracy	79.3%			
Confusion Matrix	TP= 0.95 FN= 0.05 TN=0.25 FP= 0.75			

Table 4.7: Extensions/versions of existing application & transport protocols (53 RFCs)

Results for existing A & T protocols are presented in Table 4.7. They show that backward compatibility and targeting a security extension influence success positively and negatively, respectively. Both are relatively intuitive. We expect backward compatibility to be important for all existing protocols. Conversely, we

know from experience, *e.g.,* [65], the struggles that security-motivated protocols commonly face. We also note that accuracy and TP rates are higher than for new protocols, but not so are FP rates. In other words, fewer successful protocols don't have either feature, but having those features is by itself not a guarantee of success. In other words, the features are likely necessary but far from sufficient for success.

The table also shows that even though the odds ratios obtained from the two logistic regression methods for these two features are very close, their $p - values$ are slightly different. As alluded to in the previous section, the $p - value$ from the exact method is "too conservative", and therefore, we use them as an upper bound for significance of the correlation between the corresponding feature and the success of protocols. With that in mind, the "Security" is not 95% significant, but is definitely more than 90% significant. This shows that even if the likelihood ratio test yields to close approximations of the $p - values$, it involves some errors, and using ELR we can mitigate those errors and find the accurate p-values. Also, the 90% significance is what we can achieve with this dataset, and it can be obviously improved with more samples.

In order to validate the outcomes of the stepwise regression method, we also examined the correlation of individual features with success of the protocols. The outcome shows that ELR only finds the same features as significantly correlated with success as the stepwise regression method.

4.5.2 Network Services Protocols

Network services protocols (S) have many aspects in common with A & T protocols, and therefore so do their results even if differences exist.

Table 4.8 points again to the need for new S protocols to add value to other protocols if they are to succeed, and its odds ratio and significance are even stronger than for A & T protocols. This may be because network services' primary purpose is to facilitate network usage, so that offering easier access or added functionality to other protocols is of even greater importance.

It can be seen that the $p - values$ obtained through the likelihood ratio test are accurate even though there only 28 samples in our dataset. This one more time confirms the findings of the small-data studies. Another finding is that ELR only finds the same set of features to be significantly correlated with success as the stepwise regression method.

Table 4.9 includes again backward compatibility as the one key feature for existing S protocols; one that is now present in almost all successful protocols (TP=095), though not by itself a guarantee of success (FP=0.62). Security is, however, now absent; maybe because the smaller ecosystem of network services makes the adoption of security extensions "slightly" less challenging?

The odds ratios and p-values obtained from the two logistic regression methods almost agree, even if there are slight differences (again, the exact method is too conservative in claiming significance). This is another confirmation of the studies

that claim having about 20 samples per predictor is enough for reasonable approximations of the $p - values$ obtained from the likelihood ratio test.

Additionally, upon examining whether the ELR method identifies other features than the one identified by the stepwise regression method, we realized that when "affecting network gear" is used as the single predictor variable in the ELR model, it is found to be somewhat significant ($p - value = 0.070$). However, backward compatibility is the only one that passes the 95% significance level, and when we use a model including both of these features, namely, backward compatibility and affecting network gear, their $p - values$ become 0.300 and 0.408, respectively, which are not significant at all. In other words, this shows that backward compatibility is the single most significant feature correlated with the success of existing network services protocols, and this validates the outcome of the stepwise regression method.

	Uncond. Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Gen. value for others	77.00	0.000	57.54	0.000
Accuracy	89.3%			
Confusion Matrix	TP= 0.93 FN= 0.07 TN=0.85 FP= 0.15			

Table 4.8: New network services protocols (28 RFCs)

4.5.3 Network Control Plane Protocols

Network control plane (C) protocols differ from network services protocols primarily in that they target the network as opposed to network users. So while both share close ties to the network, their success can be affected by different features, as reported in Table 4.10.

The table identifies "domain-wide value," *i.e.,* the protocol's ability to realize

	Uncond. Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Backwards Compatibility	10.80	0.014	10.10	0.049
Accuracy	68.6%			
Confusion Matrix	TP= 0.95 FN= 0.05 TN=0.38 FP= 0.62			

Table 4.9: Extensions/versions of existing network services protocols (35 RFCs)

its full value once adopted in a given domain, as an important factor in a protocol's success. This aligns with our intuition that deploying protocols that "touch" network devices is easier when their scope is limited (to a domain). The second feature, "Generating additional value for other protocols," is also consistent with the notion that network control functions that benefit other protocols should have an easier time being adopted.

The odds ratios and $p - values$ obtained by the two methods are not in agreement, even if they are close. This is due to the separation problem in our small sample size, which forced the use of Firth Bias-Adjusted regression (as the first method) instead of the unconditional logistic regression. Also, the small sample size barely meets the rule of thumb for minimum sample size of 5 to 20 samples per predictor (*i.e.,* 23 samples and 2 predictors). As a result, we have to rely on the exact method for inferences. The p-value of the two features are upper bounded by 0.056, meaning that we can draw a conclusion that both features are correlated with the success of network control plane protocols with a significance level of about 95% (94% to be accurate).

We used the ELR method to find the correlation of each individual feature with success of protocols. The outcome shows that this method only identifies the same set of features to be significantly correlated with success of protocols as the stepwise regression method does, *i.e.,* validates the outcomes of the stepwise regression method.

	Firth Bias-Adjusted Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Domain-wide value	24.64	0.008	9.78	0.056
Gen. value for other protocols	10.23	0.012	13.52	0.046
Accuracy	82.6%			
Confusion Matrix	TP= 0.7 FN= 0.30 TN=0.92 FP= 0.08			

Table 4.10: Network control plane protocols (23 RFCs)

4.5.4 Network Routing Protocols

Network routing protocols (R) include intra- and inter-domain protocols. Given the small number of RFCs involved, a single set of results is again presented, in Table 4.11, for all protocols in this category. Backward compatibility emerges again as a key feature, in part because there are few "new" protocols in this category. Another feature is "replacing another protocol," which is likely a reflection of the fact that most routing protocols have had multiple versions, with each new version replacing the previous one. Finally, "domain-wide value" is also identified as important, but with a negative impact. Its selection is somewhat ambiguous and appears driven in part by the fact that a number of intra-domain extensions did not succeed.

This may, however, change as we extend the number of RFCs under consideration, and may also be caused by "transient noise" in our labeling process[45]. In particular, most protocols associated with IPv6 have been marked as "not successful" to reflect the fact that IPv6 itself has not (yet) succeeded. This situation is, however, fast changing, as we explained in Chapter 2.

The statistics associated with these features and obtained from the two logistic regression methods demonstrate some differences. This is mainly due to the fact that, again, the odds ratios and $p - values$ of the first method are obtained using Firth Bias-Adjusted (since there is a separation problem in the data). Also again, the rule of thumb for minimum number of samples (5 to 20 samples per predictor variable) is barely met, *i.e.,* we have 32 samples and 3 predictors. This means that the statistics obtained from the exact method are more trustworthy/accurate. However, the positive finding is that the odds ratios and $p-values$ obtained from the exact method show that these features are all (positively or negatively) correlated with success of network routing protocols with a significance level of 95%.

Moreover, when comparing the ELR and stepwise regression in identifying features with significant correlation with success, we realized that the ELR method finds being a new protocol to be negatively correlated (with a significance of 0.067) with success of network routing protocols, however, this is true only when this feature is selected as the single independent variable in the model. In other words,

[45]Something that is unavoidable given that our sampling is punctual, and a protocol's success evolves over time.

upon including this feature in the above model (in addition to Backward compatibility, Replacing another protocol, and Domain wide value realization), all of their significances drop. This shows that the model identified by the stepwise regression method includes the minimal set of features that are significantly correlated with success of protocols.

	Firth Bias-Adjusted Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Backwards compatibility	99.46	0.000	28.00	0.006
Replacing another protocol	78.80	0.002	15.69	0.023
Domain-wide value	0.076	0.009	0.13	0.050
Accuracy	65.6%			
Confusion Matrix	TP= 0.86 FN= 0.14 TN=0.27 FP= 0.73			

Table 4.11: R protocols (32 RFCs)

4.5.5 Network Data Plane Protocols

Table 4.12 reports results for network data plane protocols (D) and singles out "performance improvement," as correlated with success. This aligns with our intuition

that performance is of utmost importance in the data plane, so that protocols that offer performance improvements stand a stronger chance of success.

The odds ratios obtained from the two different logistic methods are very close, however, the exact method identifies a p-value that is slightly higher than the unconditional method. The significance of this single feature is upper bounded by 0.52, which is slightly less than the 95% level of significance, and this is most probably associated with our small sample size. Additionally, the exact method only identifies "Performance" as significantly correlated with success of network data plane protocols (as does the stepwise regression method).

	Uncond. Logistic Regression		Exact Logistic Regression	
Feature	OR_u	L-R p-value	OR_c	Exact p-value
Performance Motivated	7.50	0.018	6.98	0.052
Accuracy	71.0%			
Confusion Matrix	TP= 0.50 FN= 0.50 TN=0.88 FP= 0.12			

Table 4.12: network data plane protocols (31 RFCs)

Comparing the results from the above classifications with those associated with all 230 RFCs, all new, or all existing protocols, shows that generally the accuracies, p-values, and ORs improved. Additionally, fewer features were selected for each category, making the results more interpretable. Moreover, the approximations of

the likelihood ratio test are shown to be close to the exact p-values for the small sample sizes, and therefore, trustworthy. Finally, we (roughly) validated that the outcomes of the stepwise regression method and our naive test using the ELR method always agree.

4.6 Validation

In order to further validate our findings, we constructed a new dataset from a (somewhat) different category of RFCs, namely, "Experimental". These RFCs have become mature enough to be promoted to an "Experimental" status, but the IETF has decided that they do not meet certain requirements to be promoted further to the standards track. The goal of this section is to examine this decision making process and understand whether it is consistent with the decisions to promote other RFCs to the standards track. In other words, we seek to understand whether the decision rules that decide the fate of these RFCs is similar to the one that promoted other RFCs to proposed standard and above.

Our sampling process is simple random, since we are not aware of differences in popularity of Experimental RFCs. The labeling process is similar to what we used in Section 4.3. Then we apply the models associated with each category of RFCs defined in Section 4.5.

The results vary across different types of protocols, but are mostly consistent with our intuition. In particular, the models show 100% accuracy across two cat-

egories, namely, network data plane and network control plane RFCs. In other words, we confirm that the factors associated with success or failure of "Proposed Standard" RFCs, are also correlated with the failure of "Experimental" RFCs. For other categories, however, the accuracies are not necessarily as high as the previous two categories. Specifically, for all of the other categories, the accuracies are at 65%, which is slightly lower than those of the models developed in Section 4.5. These findings show that the vetting process of the IETF is at the very least not different when deciding to promote and RFC to the standards track or not. In other words, the WG members were mostly consistent in deciding what protocols are good enough to move towards becoming a standard and what protocols are not mature enough, and need to be published with the Experimental status. Also, it shows how our models can be extended to other categories of RFCs, even if they are trained by an arguably different set.

4.7 Conclusions, Limitations & Future Works

This study investigates, using statistical tools, the correlation between a number of factors and success or failure of protocols that have reached the status of proposed standard or higher within the IETF. It shows how incorporating our engineering intuition in the process helps identify key features correlated with success or failure of protocols. The outcomes of the methodology are models that provide accuracies between 70% to 85%, with only a small number of features (between 1 to 3) for

almost all categories of RFCs. Also, the findings confirm a number of intuitive results, as well as (interesting) insight into the impact of less expected features that our engineering intuition could not afford on its own.

There are however, some limitations and shortcomings associated with our methodology and findings. In particular, the small sample sizes can potentially affect not only the accuracy of the classifiers, but also the p-value approximations requiring large samples, *e.g.,* likelihood ratio test. This is to some extent addressed by using an exact logistic regression method, and showing that the approximations are relatively accurate. However, as a future work direction, we seek to expand the dataset in cases where there are only a small number of samples and the classifier accuracy is low, *e.g.,* routing protocols, and possibly improve the statistical significance of our results.

Another potential limitation is that the accuracy of the classifier does not exceed 85%, which can be due to not including a critical factor in our study. We are open to considering additional factors in our future works given that they meet the criteria defined in Section 4.2, however, it is also likely that the adoption of a protocol depends on subtle factors that are hard, if not impossible, to measure, *e.g.,* commercial interest from companies, involvement of influential people in standardization, etc.

Moreover, we acknowledge that our study shows correlations between success or failure of a protocol and a number of features, however, correlation does not

guarantee causality. In other words, we do *not* claim that the existence or lack of the features shown in Section 4.5 *cause success or failure*. Establishing a causal relationship is much harder, and requires randomized experiments, etc., which can be investigated in future research.

Another potential shortcoming is the lack of a separate validation/test dataset, which can further confirm the outcomes. This is a direction we seek to explore in the future. In particular, by constructing a dataset from another category of RFCs, namely, those that were stopped at the Experimental stage by the IETF. This dataset not only enables us to further validate our results, but also creates a means to examine the consistency of the IETF decision making in identifying winning and losing proposals, *i.e.,* IETF has identified the RFCs in this dataset as "not good enough" to be promoted to the standards track, now this analysis allows us to compare those decisions with the ones that were indeed promoted to the standards track.

Finally, we seek to explore the robustness of our findings to other classification methods. Therefore, another future works direction can include examining non linear combination of features and using new classifiers.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

The focus of this dissertation is on understanding how and why different factors affect the adoption of a network technology. To achieve this goal, we investigate three problems, in the context of the Internet protocols. Initially, we specifically focus on IPv6, as the arguably the most important instance of protocol adoption in the Internet's history. Then learning from that case study, we extend our investigation to cover a more general Internet protocol. Our contributions involve developing methodologies for investigating the adoption of network protocols. These methodologies include developing measurement tools, models, and statistical hypotheses.

The first part of the thesis investigates the adoption of IPv6 and the path of its progress over the last two decades. It identifies the key stakeholders and decision makers that affect the evolution of IPv6 adoption. In particular, ISPs, ICPs, ITDs, and users are shown to be involved in the adoption of IPv6, either as decision makers, or as consumers of the technology. It also identifies using empirical measurements done by us and others, a 3-phase adoption pattern across different Internet stakeholders. Then it connects those phases to various factors and decisions made by the key Internet stakeholders. Finally, it develops a simple model to validate the causal relationship between the changes in those factors, and the 3 phases of IPv6 adoption. The models, and the measurement methodologies can be used in other similar technology adoption instances.

The second part of the thesis focuses on the future of IPv6 adoption, and how

different scenarios can speed up or derail its progress. The investigations involve a number of representative scenarios, and Internet stakeholders. It captures the decision making process of those Internet stakeholders and the interactions between them. The findings show that in scenarios where ISPs do not agree to offer IPv6 immediately to their users, a race to attract users can lead to a non-predictable and haphazard IPv6 adoption dynamics. However, the scenarios which include ISPs that have a minimal coordination, namely, all offer IPv6 as one of their connectivity options, the adoption picture becomes clear. Even though the latter scenarios do not offer any guarantee of IPv6 adoption, they create an environment where the impact of different factors become clear, therefore, help the Internet stakeholders to devise strategies. The study also investigates the role of address translation mechanisms, as well as the costs of translation and acquisition of extra IPv4 addresses on the adoption of IPv6. In particular, the study highlights an unintuitive and surprising finding, namely, that improvements in the performance of address translation devices can derail the progress of IPv6 adoption.

The above studies raise a question about the adoption dynamics of other protocols (or network technologies), and the impact of various factors on them. This endeavor is documented as the third part of this thesis. The study focuses on the protocols in the standards track of the Internet Engineering Task Force (IETF). First, we collect data to properly characterize those protocols using a number of features. Then, each protocol is labeled as successful or not. Finally, using statis-

tical tools, such as binary logistic regression, the correlation between the success of a protocol and the features that characterize them is investigated. This reveals a number of intuitive findings as well as some less expected ones. In particular, the findings confirm a number of anecdotal evidences, such as the positive impact of backward compatibility on the adoption of a protocol extension or version, or the negative impact of security as the motivation of a protocol. However, they also provide new insight on how generating additional value boosts the success of a new protocol, or how belonging to a specific protocol type/functionality changes the dynamics of adoption. In addition to providing the insight and formally confirming a number of anecdotal evidence about protocol adoption, the study develops a framework and a methodology on how to use statistical tools to understand the dynamics of a (general) technology adoption.

Next, we propose directions and potential future works that can improve the models and frameworks developed in this dissertation.

5.2 Potential Extensions

The research in network technology adoption is thriving along several directions. In this work we investigated the factors and features that govern the adoption of a general Internet protocol, as well as a case study of IPv6. Each of these studies have the potential for several interesting extensions as we discuss next.

In the context of IPv6 adoption, it is in particular, interesting to extend the

measurement study we started in 2009, to other services that are not necessarily provided on the Web, *e.g.,* P2P applications, mobile apps, etc. This can include developing apps that monitor other applications on a device for IPv6 usage, and record their performance. Moreover, developing models that capture the evolution of IPv6 adoption with greater details, and that have a predictive power is another interesting investigation direction. In the context of modeling the future scenarios of IPv6 adoption, the models can be extended to include strategic behaviors for different Internet stakeholders, instead of myopic ones. Also, those models can all be extended to capture a general technology adoption instance by tweaking the parameters, and interactions between the stakeholders.

There are also several extensions possible for investigating the factors or features that are correlated with the success or failure of a protocol. Generally, expanding the dataset, and providing a separate test set for further validation of the findings are useful, even if not utterly necessary given the high significance we obtain through the current dataset. Additionally, our study does not guarantee any causal relationships between the features and the outcomes, which can be extended to achieve such powers using models and/or statistical methods. Moreover, the impact of other classification methods can be investigated to obtain possible other insight into the problem. These methods include Decision Trees, Logistic Model Trees (LMT), K-Nearest-Neighbor, etc.

The findings, models, frameworks, and methodologies of this thesis contribute

not only to the networking and adoption dynamics research, but also to the research in economics, decision making, and the two sided markets. The studies also spur further research in the fields of network technology adoption, network technology design, decision making analysis, network economics etc., and provide new frameworks and tools for the researchers in these fields.

Appendices

Appendix A

Disagreement Scenario — IPv6 vs. Private IPv4

In this scenario, one ISP offers IPv6 addresses to new users, while the other relies on private IPv4 addresses. Both require translation (IPv6↔IPv4 and Private IPv4↔Public IPv4) to communicate with the public IPv4 Internet. Both types of translation equally affect connectivity quality, as measured by a common parameter, a. The greater maturity of Private to Public IPv4 translation benefits a Private IPv4 solution, since $D_4 \leq D_6$. On the flip side, IPv6 users incur translation penalties only for the fraction γ_6 of ICPs not yet IPv6 accessible.

As with the scenario of IPv6 vs. Public IPv4 (3.4.1), we describe next the decision process of users and ICPs, and how the two ISPs select their prices. For simplicity, and unlike the previous scenario where existing (public IPv4) users also had the option to adopt IPv6, we assume that only new users decide on which connectivity option to choose. Allowing existing (public IPv4) users to make such a choice requires pricing a third option (public IPv4), which adds significant complexity without qualitatively affecting the results.

A.0.1 Decision Mechanism & Solution

After the two ISPs announce prices of p_6 and p_{p4}, (new) users choose an ISP as per Eqs. (A.0.1) and (A.0.2), where $\sigma_6^{(i)}$ and $\sigma_4^{(i)}$ denote the fraction of users choosing IPv6 or private IPv4 addresses in round i, respectively. In particular, a user with quality sensitivity σ prefers IPv6 over private IPv4 if $1-p_6-\sigma a\gamma_6^{(i-1)} \geq 1-p_{p4}-\sigma a$. Note that since $\gamma_6^{(i-1)} \leq 1$, this implies that prices verify $p_{p4} \leq p_6$. Note also, that it

is possible that there exists a value σ_{NA} such that for $\sigma \geq \sigma_{\text{NA}}, 1-p_6-\sigma a\gamma_6^{(i-1)} \leq 0$, *i.e.,* users that are very sensitive to connectivity impairment will altogether opt out of connecting to the Internet. This can also arise in the previous scenario, albeit much more rarely as the availability of the public IPv4 option typically ensures that high σ users have access to a suitable alternative[46].

$$\sigma_{p4}^{(i)} = \begin{cases} \frac{(p_6-p_{p4})}{a(1-\gamma_6^{(i-1)})} & \text{if } p_{p4} > \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \\ \frac{(1-p_{p4})}{a} & \text{if } p_{p4} < \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \end{cases} \tag{A.0.1}$$

$$\sigma_6^{(i)} = \begin{cases} \frac{a(1-\gamma_6^{(i-1)})-(p_6-p_{p4})}{a(1-\gamma_6^{(i-1)})} & \text{if } p_{p4} > \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \\ & \quad \& \, p_6 < 1 - a\gamma_6^{(i-1)} \\ \frac{(1-p_6)}{a\gamma_6^{(i-1)}} - \frac{(p_6-p_{p4})}{a(1-\gamma_6^{(i-1)})} & \text{if } p_{p4} > \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \\ & \quad \& \, p_6 > 1 - a\gamma_6^{(i-1)} \\ 0 & \text{if } p_{p4} < \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \end{cases} \tag{A.0.2}$$

Once users have selected their connectivity option, ICPs proceed with their decisions as in the previous section.

$$\theta_6 = \frac{ka\sigma_6}{c_6} \tag{A.0.3}$$

[46]It arises only for combinations of large C and D_6 values, *i.e.,* very high acquisition costs for public IPv4 addresses and very high translation costs.

$$\theta_6 = \begin{cases} \min(\max(\frac{ka(1-\gamma_6^{(i-1)})-k(p_6-p_{p4})}{c_6(1-\gamma_6^{(i-1)})}, 1-\gamma_6^{(i-1)}), 1) \\ \text{if } p_{p4} > \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \,\&\, p_6 < 1 - a\gamma_6^{(i-1)} \\ \min(\max(\frac{k(1-p_6)}{c_6\gamma_6^{(i-1)}} - \frac{k(p_6-p_{p4})}{c_6(1-\gamma_6^{(i-1)})}, 1-\gamma_6^{(i-1)}), 1) \\ \text{if } p_{p4} > \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \,\&\, p_6 > 1 - a\gamma_6^{(i-1)} \\ \min(\max(0, 1-\gamma_6^{(i-1)}), 1) \\ \text{if } p_{p4} < \frac{p_6-1+\gamma_6^{(i-1)}}{\gamma_6^{(i-1)}} \end{cases} \tag{A.0.4}$$

$$p_{p4} = \underset{p_{p4}}{\text{argmax}}\{(1+i\delta)(\sigma_{p4})p_{p4} - D_4\sigma_{p4}(1+i\delta)\} \tag{A.0.5}$$

$$p_6 = \underset{p_6}{\text{argmax}}\{(1+i\delta)\sigma_6 p_6 - D_6\sigma_6(1+i\delta)\gamma_6^{(i)}\} \tag{A.0.6}$$

Solving Eq. (A.0.5) requires an optimization with respect to the conditions of Eq. (A.0.4), however, those conditions are cumbersome, hence, we resort to numerical analysis.

The Impact of Disagreement: Fig. A.1 offers a perspective similar to that of Fig. 3.2, and reports the outcome of the ISPs' price selection process for a range of configurations. There are some differences, *e.g.,* migration to IPv6 arises rarely, but there is nevertheless a broad range of parameters for which cycle are present. As argued earlier, this makes devising pricing strategies difficult and is likely to contribute continued uncertainty in deciding how to settle on a migration strategy.

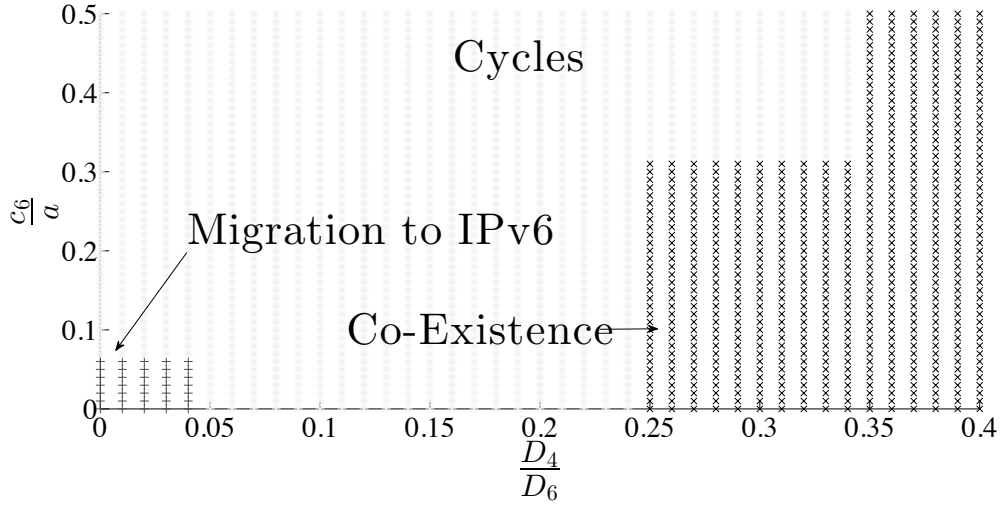


Figure A.1: IPv6 vs. private IPv4 competition

Appendix B

Robustness Tests — Figures

Here, we present a set of figures similar to Fig. 3.4 for the robustness tests of Section 3.6. These figures show that despite slight quantitative differences introduced by various assumptions in the model, the outcomes are qualitatively similar to the original solution of Section 3.4.

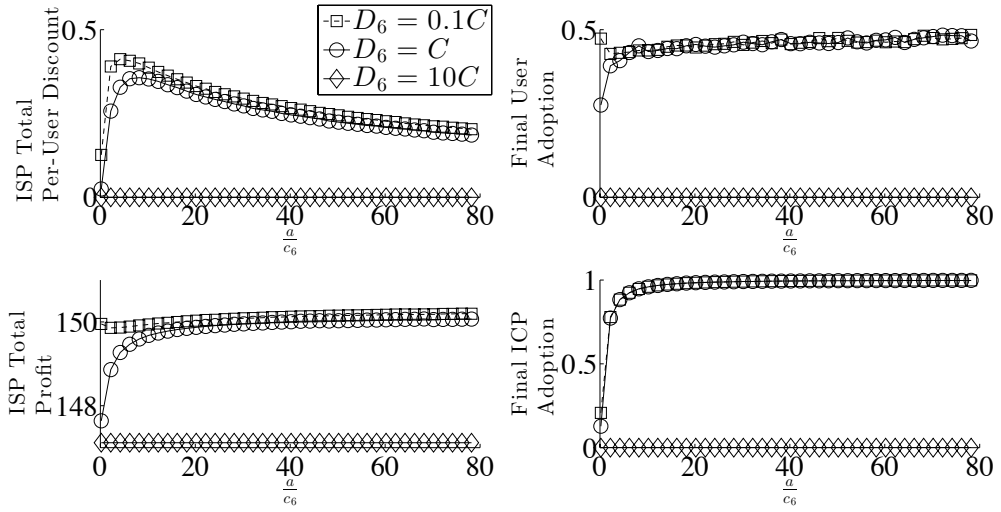


Figure B.1: Total profit, discount & adoption levels for small C — single-modal β

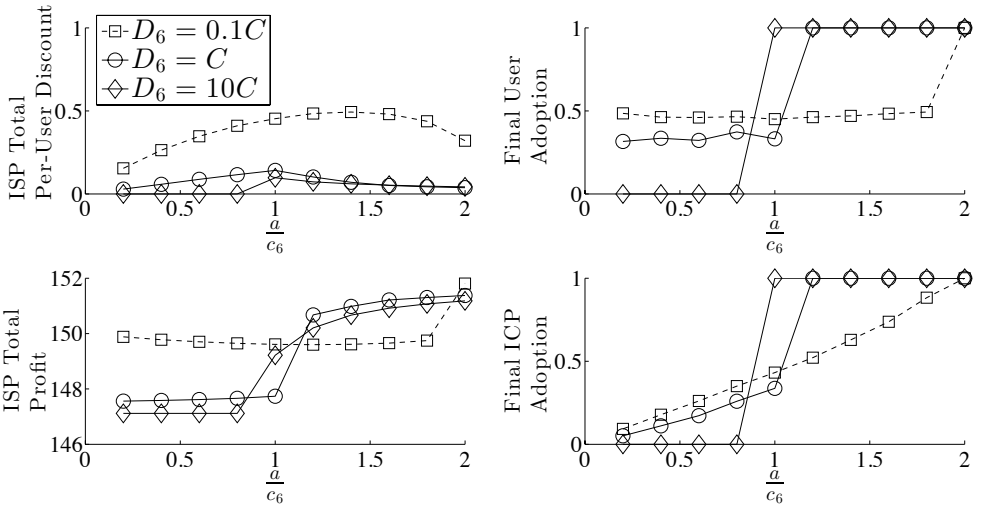


Figure B.2: Total profit, discount & adoption levels for small C — single-modal σ

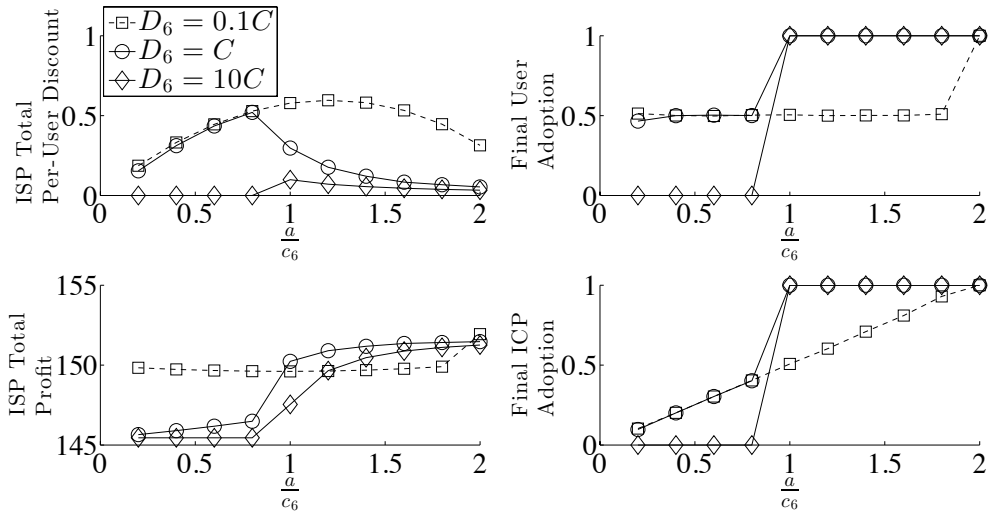


Figure B.3: Total profit, discount & adoption levels for small C — Linear IPv4 acquisition cost

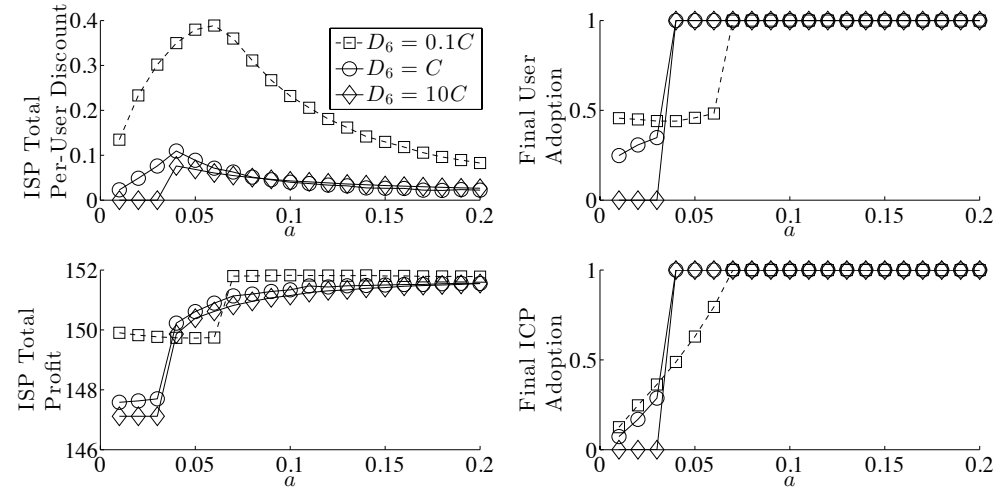


Figure B.4: Total profit, discount & adoption levels for small C — Decreasing c_6

Appendix C

ICPs Lag Behind Users & ISPs

Here, we investigate a model in which the timing of ICPs' decisions is different from that of the ISPs and users. In other words, ICPs less frequently make decisions compared to the ISPs and users. From a modeling standpoint, we can remove the decision making of ICPs, and update the fraction of IPv6 accessible ICPs infrequently. In this setting, we can analytically find the Nash Equilibrium of the game played by ISPs (in a disagreement scenario).

This setting is analytically similar to the one where ICPs make decisions *after* each price announcement, but the timing of their decision making remains unknown to ISPs. The model formulation remains the same (as the original model of Section 3.3), except for the IPv6 ISP's utility function, which instead of $\gamma_6^{(i)}$ uses the previous fraction of IPv4-only ICPs ($\gamma_6^{(i-1)}$). In other words, instead of predicting what ICPs will do after announcement of prices, the IPv6 ISP uses the known number of IPv6-accessible ICPs. Next, we provide the analytical solution to this model.

$$\sigma = \frac{p_4 - p_6}{a\gamma_6^{(i-1)}}$$

$$\Rightarrow n_6 = s\sigma \ , \ n_4 = s(1-\sigma)$$

$$f(p_4) = n_4 p_4 - C(n_4 - 1)_+^2$$

$$g(p_6) = n_6 p_6 - D_6 n_6 \gamma_6^{(i-1)}$$

$$f'(p_4) = 0$$

$$\Rightarrow s - \frac{2s}{a\gamma_6^{(i-1)}} p_4 + \frac{s}{a\gamma_6^{(i-1)}} p_6 +$$
$$\frac{2Cs^2}{a\gamma_6^{(i-1)}} - \frac{2Cs}{a\gamma_6^{(i-1)}} - \frac{2Cs^2}{(a\gamma_6^{(i-1)})^2} p_4 + \frac{2Cs^2}{(a\gamma_6^{(i-1)})^2} p_6 = 0$$
$$\Rightarrow p_4^* = \frac{1}{2 + \frac{2Cs}{a\gamma_6^{(i-1)}}} ((1 + \frac{2Cs}{a\gamma_6^{(i-1)}}) p_6 + a\gamma_6^{(i-1)} + 2Cs - 2C)$$

$$g'(p_6) = 0$$

$$\Rightarrow \frac{s}{a\gamma_6^{(i-1)}} p_4 - \frac{2s}{a\gamma_6^{(i-1)}} p_6 + \frac{D_6 s}{a} = 0$$
$$\Rightarrow p_6^* = \frac{p_4}{2} + \frac{D_6 \gamma_6^{(i-1)}}{2}$$

where s is the total user population at epoch i, p_4^* and p_6^* are the best response prices of the IPv4 and IPv6 ISPs, respectively. Depending on the values of these prices, the solution space is divided into three regions:

Region 1:

$$0 \leq p_4 * - p_6 * \leq a\gamma_6^{(i-1)} \;\; \& \;\; n_4 > 1$$

$$a > D_6 \;\; \& \;\; s \geq \frac{3a}{2a + D_6}$$

OR

$$a < D_6 \;\; \& \;\; s \geq 1 + \frac{\gamma_6^{(i-1)}(D_6 - a)}{2C}$$

In this region, the Nash Equilibrium can be found as follows:

$$p_4^* = \Big(aD_6(\gamma_6^{(i-1)})^2 + 2D_6\gamma_6^{(i-1)}Cs +$$
$$2(a\gamma_6^{(i-1)})^2 + 4a\gamma_6^{(i-1)}Cs - 4a\gamma_6^{(i-1)}C\Big) / \Big(a\gamma_6^{(i-1)} + 2Cs\Big)$$

$$p_6^* = \Big(aD_6(\gamma_6^{(i-1)})^2 + 2D_6\gamma_6^{(i-1)}Cs +$$
$$2(a\gamma_6^{(i-1)})^2 + 4a\gamma_6^{(i-1)}Cs - 4a\gamma_6^{(i-1)}C\Big) / \Big(6a\gamma_6^{(i-1)} + 4Cs\Big)$$
$$+ \frac{D_6\gamma_6^{(i-1)}}{2}$$

Region 2:

$$0 \leq p_4 * - p_6 * \leq a\gamma_6^{(i-1)} \;\; \& \;\; n_4 \leq 1$$

$$a > D_6 \;\; \& \;\; 1 < s < \frac{3a}{2a + D_6}$$

The Nash Equilibrium in this region is obtained as follows:

$$p_4^* = \frac{p_6}{2} + \frac{a\gamma_6^{(i-1)}}{2}$$
$$p_6^* = \frac{p_4}{2} + \frac{D_6\gamma_6^{(i-1)}}{2}$$
$$\Rightarrow p_4^* = \frac{D_6\gamma_6^{(i-1)}}{3} + \frac{2a\gamma_6^{(i-1)}}{3}$$

Region 3:

$$a < D_6 \ \ \& \ \ 1 < s < 1 + \frac{\gamma_6^{(i-1)}(D_6 - a)}{2C}$$

In this region, there can be two Nash Equilibria:

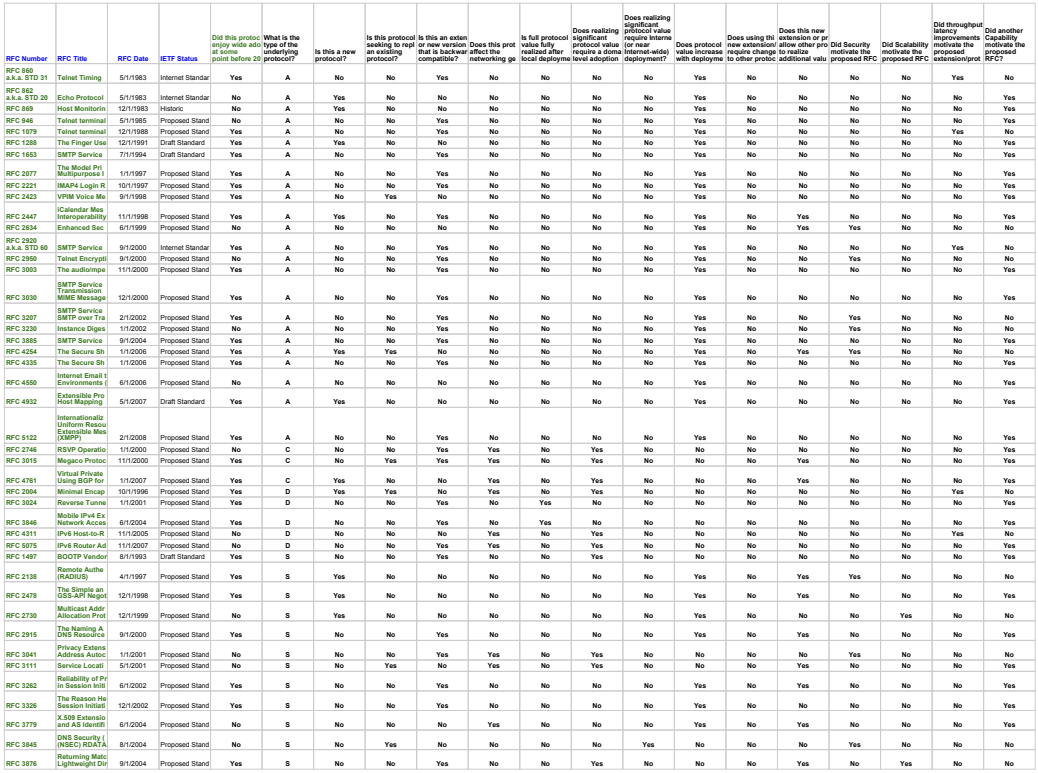
$$p_4^* = p_6^* \ \text{ No IPv6 Users}$$

OR

$$p_4^* - p_6^* \geq a\gamma_6^{(i-1)} \ \text{ No IPv4 Users}$$

Appendix D

RFC Labels and Characteristics



RFC Number	RFC Title	RFC Date	IETF Status	Did this protoc enjoy wide ado at some point before 20	What is the type of the underlying protocol?	Is this a new protocol?	Is this protocol seeking to repl an existing protocol?	Is this an exten or new version that is backwar compatible?	Does this prot affect the networking ge	Is full protocol value fully realized after local deployme	Does realizing significant protocol value require a doma level adoption	Does realizing significant protocol value require Interne (or near internet-wide) deployment?	Does protocol value increase with deployme	Does using thi new extension require change to other protoc	Does this new extension or pr allow other pro to realize additional valu	Did Security motivate the proposed RFC	Did Scalability motivate the proposed RFC	Did throughput latency improvements motivate the proposed extension/prot	Did another Capability motivate the proposed RFC?
RFC 860 a.k.a. STD 31	Telnet Timing	5/1/1983	Internet Standar	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 862 a.k.a. STD 20	Echo Protocol	5/1/1983	Internet Standar	No	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 869	Host Monitorin	12/1/1983	Historic	No	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 946	Telnet terminal	5/1/1985	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1079	Telnet terminal	12/1/1988	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 1288	The Finger Use	12/1/1991	Draft Standard	Yes	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1653	SMTP Service	7/1/1994	Draft Standard	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2077	The Model Pri Multipurpose I	1/1/1997	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2221	IMAP4 Login R	10/1/1997	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2423	VPIM Voice Me	9/1/1998	Proposed Stand	Yes	A	No	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2447	iCalendar Mes Interoperability	11/1/1998	Proposed Stand	Yes	A	Yes	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 2634	Enhanced Sec	6/1/1999	Proposed Stand	No	A	No	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 2920 a.k.a. STD 60	SMTP Service	9/1/2000	Internet Standar	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 2950	Telnet Encrypti	9/1/2000	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 3003	The audio/mpe	11/1/2000	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3030	SMTP Service Transmission MIME Messaga	12/1/2000	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3207	SMTP Service SMTP over Tra	2/1/2002	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 3230	Instance Diges	1/1/2002	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 3885	SMTP Service	9/1/2004	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4254	The Secure Sh	1/1/2006	Proposed Stand	Yes	A	Yes	Yes	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4335	The Secure Sh	1/1/2006	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4550	Internet Email t Environments (	6/1/2006	Proposed Stand	No	A	No	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4932	Extensible Pro Host Mapping	5/1/2007	Draft Standard	Yes	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 5122	Internationaliz Uniform Resou Extensible Mes (XMPP)	2/1/2008	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2746	RSVP Operatio	1/1/2000	Proposed Stand	No	C	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3015	Megaco Protoc	11/1/2000	Proposed Stand	Yes	C	No	Yes	Yes	No	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 4761	Virtual Private Using BGP for	1/1/2007	Proposed Stand	Yes	C	Yes	No	No	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2004	Minimal Encap	10/1/1996	Proposed Stand	Yes	D	Yes	Yes	No	Yes	No	Yes	No	No	No	No	No	No	Yes	No
RFC 3024	Reverse Tunne	1/1/2001	Proposed Stand	Yes	D	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No	Yes
RFC 3846	Mobile IPv4 Ex Network Acces	6/1/2004	Proposed Stand	Yes	D	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No	Yes
RFC 4311	IPv6 Host-to-R	11/1/2005	Proposed Stand	No	D	No	No	No	Yes	No	Yes	No	No	No	No	No	No	Yes	No
RFC 5075	IPv6 Router Ad	11/1/2007	Proposed Stand	No	D	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1497	BOOTP Vendor	8/1/1993	Draft Standard	Yes	S	No	No	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2138	Remote Authe (RADIUS)	4/1/1997	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 2478	The Simple an GSS-API Negot	12/1/1998	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	Yes
RFC 2730	Multicast Addr Allocation Prot	12/1/1999	Proposed Stand	No	S	Yes	No	No	No	No	No	No	Yes	No	No	No	Yes	No	No
RFC 2915	The Naming A DNS Resourca	9/1/2000	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3041	Privacy Extens Address Autoc	1/1/2001	Proposed Stand	No	S	No	No	Yes	Yes	No	Yes	No	No	No	No	Yes	No	No	No
RFC 3111	Service Locati	5/1/2001	Proposed Stand	No	S	No	Yes	No	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 3262	Reliability of Pr in Session Initi	6/1/2002	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3326	The Reason He Session Initiati	12/1/2002	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3779	X.509 Extensio and AS Identifi	6/1/2004	Proposed Stand	No	S	No	No	No	Yes	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3845	DNS Security ((NSEC) RDATA	8/1/2004	Proposed Stand	No	S	No	Yes	No	No	No	No	Yes	No	No	No	Yes	No	No	No
RFC 3876	Returning Matc Lightweight Dir	9/1/2004	Proposed Stand	Yes	S	No	No	Yes	No	No	Yes	No	No	No	Yes	No	Yes	No	No

Figure D.1: 230 RFCs - Labels and Characteristics

RFC 3911	The Session In "Join" Header	10/1/2004	Proposed Stand	No	S	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3912	WHOIS Protoc	9/1/2004	Draft Standard	Yes	S	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4032	Update to the Preconditions	3/1/2005	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4537	Kerberos Crypt	6/1/2006	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4568	Session Descri Security Descr	7/1/2006	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4574	The Session D Label Attribute	8/1/2006	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 4680	TLS Handshak	10/1/2006	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 5055	Server-Based	12/1/2007	Proposed Stand	No	S	Yes	Yes	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 5238	Datagram Tran over the Datag	5/1/2008	Proposed Stand	No	S	No	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 1692	Transport Multi	8/1/1994	Historic (change Proposed Stand March 2006)	No	T	Yes	No	No	No	No	No	No	Yes	Yes	No	No	No	Yes	No
RFC 2032	RTP Payload F	10/1/1996	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2793	RTP Payload f	5/1/2000	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3390	Increasing TCP	10/1/2002	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 3517	A Conservativ (SACK)-based	4/1/2003	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 4960	Stream Control	9/1/2007	Proposed Stand	No	T	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 5215	RTP Payload F	8/1/2008	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	Yes	No	No	No	No	Yes
RFC 5285	A General Mec	7/1/2008	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 1397	Default Route BGP3 Version	1/1/1993	Historic (change Proposed Stand March 2006)	Yes	R	No	No	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No	No
RFC 3065	Autonomous S	2/1/2001	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	Yes	No	No
RFC 4486	Subcodes for	4/1/2006	Proposed Stand	Yes	R	No	No	Yes	Yes	Yes	No	No	No	No	No	No	No	No	Yes
RFC 1997	BGP Communi	8/1/1996	Proposed Stand	Yes	R	No	No	Yes	Yes	No	No	No	Yes	No	No	No	Yes	No	Yes
RFC 1519	Classless Inter an Address As Aggregation St	9/1/1993	Proposed Stand	Yes	R	No	Yes	No	Yes	No	No	No	Yes	Yes	No	No	Yes	No	No
RFC 2439	BGP Route Fla	11/1/1998	Proposed Stand	Yes	R	No	No	Yes	Yes	No	No	No	Yes	No	No	No	Yes	No	No
RFC 4238	Voice Message	10/1/2005	Proposed Stand	No	R	Yes	No	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2918	Route Refresh	9/1/2000	Proposed Stand	Yes	R	No	No	Yes	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1388	RIP Version 2	1/1/1993	Proposed Stand	Yes	R	No	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	No	No	Yes
RFC 4203	OSPF Extensio Multi-Protocol	10/1/2005	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 4456	BGP Route Ref Full Mesh Inter	4/1/2006	Draft Standard	Yes	R	No	No	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No	No
RFC 4781	Graceful Resta	1/1/2007	Proposed Stand	Yes	R	No	No	Yes	Yes	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 5089	IS-IS Protocol Element (PCE)	1/1/2008	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5308	Routing IPv6 w	10/1/2008	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 5329	Traffic Enginee	9/1/2008	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2453	RIP Version 2	11/1/1998	Internet Standar	Yes	R	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1587	The OSPF NSS	3/1/1994	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5304	IS-IS Cryptogra	10/1/2008	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	No	Yes	No	No	No
RFC 3623	Graceful OSPF	11/1/2003	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 4915	Multi-Topology	6/1/2007	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5301	Dynamic Hostn Mechanism for	10/1/2008	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3107	Carrying Label	5/1/2001	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	Yes	No	Yes
RFC 1184	Telnet Linemo	10/1/1990	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1413	Identification P	2/1/1993	Proposed Stand	Yes	A	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 1647	TN3270 Enhan	7/1/1994	Proposed Stand	Yes	A	No	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1730	Internet Messa	12/1/1994	Proposed Stand	Yes	A	No	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1782	TFTP Option E	3/1/1995	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 1869	SMTP Service	11/1/1995	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1893	Enhanced Mail	1/1/1996	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1939	Post Office Pro	5/1/1996	Proposed Stand	Yes	A	No	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2045	Multipurpose I Part One: Form	11/1/1996	Draft Standard	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2068	Hypertext Tran HTTP/1.1	1/1/1997	Proposed Stand	Yes	A	No	Yes	Yes	No	No	No	No	Yes	No	Yes	Yes	No	No	Yes
RFC 2156	MIXER (Mime I Mapping betwe	1/1/1998	Proposed Stand	No	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2244	ACAP -- Applic	11/1/1997	Proposed Stand	No	A	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2371	Transaction Int	7/1/1998	Proposed Stand	No	A	No	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2389	Feature negoti	8/1/1998	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2630	Cryptographic	6/1/1999	Proposed Stand	No	A	No	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No

Figure D.2: 230 RFCs - Labels and Characteristics (Continued)

RFC 2821	Simple Mail Tr	4/1/2001	Proposed Stand	Yes	A	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3335	MIME-based S Business Data over the Intern	9/1/2002	Proposed Stand	No	A	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 3530	Network File S	4/1/2003	Proposed Stand	Yes	A	No	Yes	Yes	No	No	No	No	Yes	No	No	Yes	No	No	Yes
RFC 3730	Extensible Pro	3/1/2004	Proposed Stand	Yes	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3887	Message Track	9/1/2004	Proposed Stand	No	A	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3920	Extensible Mes	10/1/2004	Proposed Stand	Yes	A	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3977	Network News	10/1/2006	Proposed Stand	Yes	A	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	No	No	No	Yes
RFC 4239	Internet Voice	11/1/2005	Proposed Stand	No	A	No	No	No	No	No	No	No	Yes	Yes	Yes	No	No	No	Yes
RFC 4314	IMAP4 Access	12/1/2005	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 4975	The Message S	9/1/2007	Proposed Stand	Yes	A	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 5023	The Atom Publ	10/1/2007	Proposed Stand	No	A	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1157	Simple Networ	5/1/1990	Proposed Stand	Yes	C	Yes	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1256	ICMP Router Di	9/1/1991	Proposed Stand	Yes	C	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1352	SNMP Security	7/1/1992	Historic (Chang from Proposed Standard in 200	No	C	No	No	Yes	Yes	No	Yes	No	No	No	No	Yes	No	No	No
RFC 2205	Resource ReS Functional Spe	9/1/1997	Proposed Stand	No	C	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No	No	Yes
RFC 2741	Agent Extensib	1/1/2000	Draft Standard	Yes	C	Yes	No	No	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2748	The COPS (Co	1/1/2000	Proposed Stand	No	C	Yes	No	No	Yes	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 3292	General Switch	6/1/2002	Proposed Stand	No	C	No	Yes	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 3376	Internet Group	10/1/2002	Proposed Stand	No	C	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3525	Gateway Contr	6/1/2003	Historic (change from Proposed Standard)	Yes	C	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 4601	Protocol Indep Protocol Speci	8/1/2006	Proposed Stand	No	C	Yes	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 4874	Exclude Route Protocol-Traffi	4/1/2007	Proposed Stand	Yes	C	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 1209	The Transmiss	3/1/1991	Internet Standar	Yes	D	No	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 1237	Guidelines for	7/1/1991	Proposed Stand	No	D	Yes	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No	No	Yes
RFC 1349	Type of Servic	7/1/1992	Proposed Stand	No	D	No	No	Yes	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1420	SNMP over IPX	3/1/1993	Proposed Stand	Yes	D	No	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1575	An Echo Funct	2/1/1994	Draft Standard	No	D	No	No	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1825	Security Archit	8/1/1995	Proposed Stand	No	D	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No
RFC 1883	Internet Protoc	12/1/1995	Proposed Stand	No	D	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	No
RFC 2002	IP Mobility Sup	1/1/2002	Proposed Stand	No	D	No	No	Yes	No	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 2474	Definition of th in the IPv4 and	12/1/1998	Proposed Stand	Yes	D	No	No	Yes	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No
RFC 3031	Multiprotocol L	1/1/2001	Proposed Stand	Yes	D	Yes	No	No	Yes	No	Yes	No	No	Yes	Yes	No	No	Yes	No
RFC 3095	RObust Header and four profil	7/1/2001	Proposed Stand	Yes	D	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 3775	Mobility Suppo	6/1/2004	Proposed Stand	No	D	No	Yes	Yes	No	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 3931	Layer Two Tun	3/1/2005	Proposed Stand	Yes	D	No	Yes	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3963	Network Mobili	1/1/2005	Proposed Stand	No	D	No	No	Yes	No	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 4433	Mobile IPv4 Dy	3/1/2006	Proposed Stand	Yes	D	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No	Yes
RFC 4945	The Internet IP IKEv2, and PKI	8/1/2007	Proposed Stand	No	D	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 1971	IPv6 Stateless	8/1/1996	Proposed Stand	No	S	Yes	No	No	No	Yes	No	No	No	No	No	No	No	No	Yes
RFC 2131	Dynamic Host	3/1/1997	Proposed Stand	Yes	S	No	Yes	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2136	Dynamic Updat	4/1/1997	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	Yes	No	No
RFC 2165	Service Locati	6/1/1997	Proposed Stand	Yes	S	Yes	No	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2651	The Architectu	8/1/1999	Proposed Stand	No	S	No	Yes	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 2848	The PINT Servi SDP for IP Acc	6/1/2000	Proposed Stand	No	S	No	No	Yes	No	No	No	No	Yes	Yes	Yes	No	No	No	Yes
RFC 2911	Internet Printin	9/1/2000	Proposed Stand	Yes	S	No	Yes	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3261	SIP: Session In	6/1/2002	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3340	The Applicatio	7/1/2002	Historic (change from Proposed Standard)	No	S	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3456	Dynamic Host Configuration	1/1/2003	Proposed Stand	No	S	No	No	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3588	Diameter Base	9/1/2003	Proposed Stand	Yes	S	Yes	Yes	No	No	No	Yes	No	No	No	Yes	Yes	No	No	No
RFC 3596	DNS Extension	10/1/2003	Draft Standard	No	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3927	Dynamic Confi	5/1/2005	Proposed Stand	Yes	S	Yes	No	No	No	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 3959	The Early Sess Protocol (SIP)	12/1/2004	Proposed Stand	No	S	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4120	The Kerberos	7/1/2005	Proposed Stand	Yes	S	No	Yes	Yes	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4346	The Transport	4/1/2006	Proposed Stand	Yes	S	No	Yes	Yes	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4422	Simple Authen	6/1/2006	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No

Figure D.3: 230 RFCs - Labels and Characteristics (Continued)

RFC 4430	Kerberized Inte	3/1/2006	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 4555	IKEv2 Mobility	6/1/2006	Proposed Stand	No	S	No	No	Yes	No	Yes	No	No	No	No	Yes	Yes	No	No	No
RFC 4582	The Binary Flo	11/1/2006	Proposed Stand	Yes	S	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 5189	Middlebox Co	3/1/2008	Proposed Stand	No	S	Yes	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5191	Protocol for Ca (PANA)	5/1/2008	Proposed Stand	No	S	Yes	No	No	No	No	Yes	No	No	No	Yes	Yes	No	No	No
RFC 5216	The EAP-TLS	3/1/2008	Proposed Stand	Yes	S	Yes	No	No	No	No	Yes	No	No	No	Yes	Yes	No	No	No
RFC 5389	Session Traver	10/1/2008	Proposed Stand	Yes	S	Yes	No	No	No	Yes	No	No	No	No	Yes	No	No	No	Yes
RFC 1323	TCP Extension	5/1/1992	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	Yes	No
RFC 2581	TCP Congestio	4/1/1999	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 2883	An Extension t (SACK) Option	7/1/2000	Proposed Stand	Yes	T	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 3550	RTP: A Transp	7/1/2003	Proposed Stand	Yes	T	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3720	Internet Small	4/1/2004	Proposed Stand	Yes	T	Yes	Yes	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 4340	Datagram Con	3/1/2006	Proposed Stand	No	T	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2080	RIPng for IPv6	1/1/1997	Proposed Stand	No	R	No	Yes	No	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2328	OSPF Version	4/1/1998	Internet Standar	Yes	R	No	Yes	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5088	OSPF Protocol Element (PCE)	1/1/2008	Proposed Stand	No	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 5305	IS-IS Extension	10/1/2008	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 1478	An Architectur	6/1/1993	Historic (change from Proposed Standard)	No	R	Yes	No	No	Yes	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes
RFC 3219	Telephony Rou	1/1/2002	Proposed Stand	No	R	Yes	No	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4271	A Border Gate	1/1/2006	Draft Standard	Yes	R	No	Yes	Yes	Yes	No	No	Yes	No	No	No	No	No	No	Yes
RFC 1577	Classical IP an	1/1/1994	Proposed Stand	No	D	Yes	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 1731	IMAP4 Authent	12/1/1994	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 1835	Architecture of	8/1/1995	Historic (Proposed Stan	No	A	No	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 1905	Protocol Opera Network Mana	1/1/1996	Draft Standard	No	C	No	Yes	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2003	IP Encapsulati	10/1/1996	Proposed Stand	Yes	D	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	Yes
RFC 2163	Using the Inter Conformant GI	1/1/1998	Proposed Stand	No	A	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2207	RSVP Extensio	9/1/1997	Proposed Stand	No	C	No	No	No	Yes	No	No	No	Yes	No	No	Yes	No	No	No
RFC 2234	Augmented BN	11/1/1997	Proposed Stand	Yes	A	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 2236	Internet Group	11/1/1997	Proposed Stand	No	C	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2289	A One-Time Pa	2/1/1998	Internet Standar	Yes	A	Yes	Yes	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 2453	RIP Version 2	11/1/1998	Internet Standar	Yes	R	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2784	Generic Routin	3/1/2000	Proposed Stand	No	D	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 2486	The Network A	1/1/1999	Proposed Stand	Yes	S	Yes	No	No	Yes	No	Yes	No	No	No	Yes	Yes	No	No	No
RFC 2557	MIME Encapsu such as HTML	3/1/1999	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2608	Service Locati	6/1/1999	Proposed Stand	Yes	S	No	Yes	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 2625	IP and ARP ov	6/1/1999	Proposed Stand	No	D	No	No	No	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 2652	MIME Object D Common Index	8/1/1999	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2750	RSVP Extensio	1/1/2000	Proposed Stand	No	C	No	No	Yes	Yes	No	Yes	No	No	No	No	No	Yes	No	No
RFC 2776	Multicast-Scop Protocol (MZA	2/1/2000	Historic (Proposed Stan	No	S	Yes	No	No	Yes	No	Yes	No	No	No	No	Yes	No	No	No
RFC 2893	Transition Mec Hosts and Rou	8/1/2000	Proposed Stand	No	D	No	No	Yes	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 2913	MIME Content	9/1/2000	Proposed Stand	No	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3007	Secure Domain Dynamic Updat	11/1/2000	Proposed Stand	No	S	No	Yes	No	Yes	No	No	Yes	No	No	No	Yes	No	No	No
RFC 3080	The Blocks Ext	3/1/2001	Proposed Stand	No	S	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3203	DHCP reconfig	12/1/2001	Proposed Stand	Yes	S	No	No	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3380	Internet Printin Job and Printe	9/1/2002	Proposed Stand	Yes	S	No	No	Yes	No	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3403	Dynamic Deleg Part Three: The Domain Na	10/1/2002	Proposed Stand	No	S	Yes	No	No	Yes	No	No	No	Yes	Yes	No	No	Yes	No	No
RFC 3648	Web Distributa Ordered Collec	12/1/2003	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3659	Extensions to	3/1/2007	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3910	The SPIRITS (S Internet Servic	10/1/2004	Proposed Stand	No	S	Yes	No	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 3965	A Simple Mode	12/1/2004	Draft Standard	Yes	A	Yes	No	No	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 3981	IRIS: The Inter Core Protocol	1/1/2005	Proposed Stand	No	S	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4037	Open Pluggabl Protocol (OCP)	3/1/2005	Proposed Stand	No	S	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	Yes

Figure D.4: 230 RFCs - Labels and Characteristics (Continued)

RFC 4124	Protocol Exten MPLS Traffic E	6/1/2005	Proposed Stand	Yes	D	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	Yes	No
RFC 4287	The Atom Syn	12/1/2005	Proposed Stand	No	A	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4302	IP Authenticati	12/1/2005	Proposed Stand	No	D	Yes	Yes	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No
RFC 4344	The Secure Sh Modes	1/1/2006	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 4448	Encapsulation over MPLS Net	4/1/2006	Proposed Stand	Yes	D	No	No	No	Yes	No	Yes	No	No	No	No	No	No	Yes	No
RFC 4511	Lightweight Dir The Protocol	6/1/2006	Proposed Stand	Yes	S	Yes	Yes	No	No	No	No	No	Yes	No	Yes	No	Yes	No	No
RFC 4535	GSAKMP: Gro Management P	6/1/2006	Proposed Stand	No	S	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 4607	Source-Specifi	8/1/2006	Proposed Stand	No	C	No	No	Yes	Yes	No	No	No	Yes	No	No	No	Yes	No	No
RFC 4644	Network News Extension for	10/1/2006	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	Yes
RFC 4719	Transport of El Tunneling Prot	11/1/2006	Proposed Stand	Yes	C	No	No	Yes	Yes	No	Yes	No	No	No	Yes	No	No	No	Yes
RFC 4741	NETCONF Con	12/1/2006	Proposed Stand	Yes	C	Yes	No	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	No
RFC 4762	Virtual Private Distribution Pr	1/1/2007	Proposed Stand	Yes	C	Yes	No	No	Yes	No	Yes	No	No	No	No	No	Yes	No	No
RFC 4866	Enhanced Rou	5/1/2007	Proposed Stand	No	D	No	No	Yes	Yes	Yes	No	No	No	No	No	Yes	No	Yes	No
RFC 4978	The IMAP COM	8/1/2007	Proposed Stand	Yes	A	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	No
RFC 5018	Connection Es Control Protoc	9/1/2007	Proposed Stand	Yes	S	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	Yes
RFC 5082	The Generalize	10/1/2007	Proposed Stand	Yes	D	Yes	No	No	Yes	No	No	No	Yes	No	No	Yes	No	No	No
RFC 5090	RADIUS Exten	2/1/2008	Proposed Stand	No	S	No	No	No	No	No	No	No	Yes	No	Yes	Yes	No	No	No
RFC 5246	The Transport Version 1.2	8/1/2008	Proposed Stand	Yes	T	No	Yes	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No
RFC 5316	ISIS Extension System (AS) M	12/1/2008	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	Yes	No
RFC 5348	TCP Friendly R Specification	9/1/2008	Proposed Stand	No	T	Yes	No	No	No	No	No	No	Yes	Yes	No	No	No	Yes	No
RFC 5357	A Two-Way Act	10/1/2008	Proposed Stand	No	C	Yes	No	No	Yes	No	No	No	Yes	No	No	No	No	No	Yes
RFC 5303	Three-Way Han Adjacencies	10/1/2008	Proposed Stand	Yes	R	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	Yes
RFC 3748	Extensible Aut	6/1/2004	Proposed Stand	Yes	S	Yes	No	No	No	No	Yes	No	No	No	Yes	Yes	No	No	No
RFC 5386	Better-Than-No Mode of IPsec	11/1/2008	Proposed Stand	No	S	No	No	Yes	No	No	No	No	Yes	No	No	Yes	No	No	No

Figure D.5: 230 RFCs - Labels and Characteristics (Continued)

Bibliography

[1] Available at `https://en.wikipedia.org/wiki/Logistic_regression`.

[2] A. Ozment and S.E. Schechter. Bootstrapping the adoption of Internet security protocols. In *Proc. WEIS*, 2006.

[3] S. Alcock, R. Nelson, and D. Miles. Investigating the impact of service provider NAT on residential broadband users. In *Proc. IEEE INFOCOM*, San Diego, California, 2010.

[4] M. Armstrong. Competition in two-sided markets. *The RAND J. Econ.*, 37(3), 2006.

[5] M. Bagnulo, P. Matthews, and I. van Beijnum. Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers. RFC 6146 (Proposed Standard), Apr. 2011.

[6] V. Bajpai and J. Schönwälder. IPv4 versus IPv6-Who connects faster? In *Proc. IFIP Networking Conference, Toulouse, France*, 2015.

[7] M. Bartlett. Approximate confidence intervals: Iii. a bias correction. *Biometrika*, pages 201–204, 1955.

[8] R. Beverly, M. Luckie, L. Mosley, and K. Claffy. Measuring and Characterizing IPv6 Router Availability. In *Passive and Active Measurement*. Springer, 2015.

[9] J. Brutlag. Speed matters for Google web search, 2009. (Available at `http://goo.gl/UfxXOT`).

[10] K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, 2002.

[11] A historical view of the AS core. (Available at `http://goo.gl/OhqWNM`).

[12] State of IPv6 in China. (Available at `http://goo.gl/sXG6gO`).

[13] IPv6 Deployment Best Practice by China Telecom. (Available at `http://goo.gl/8OewsD`).

[14] About Carrier Grade NAT (CGN). (Available at `http://goo.gl/S2gF5e`).

[15] K. Cho, M. Luckie, and B. Huffaker. Identifying IPv6 network problems in the dual-stack world. In *Proc. ACM SIGCOMM NetT Workshop*, 2004.

[16] The role of government in IPv6 adoption. Cisco White Paper, 2010. (Available at `http://goo.gl/iiSJAS`).

[17] k. claffy. Tracking IPv6 evolution: Data we have and data we need. *ACM SIGCOMM Comput. Comm. Rev.*, 41(3), 2011.

[18] L. Colitti, S. Gunderson, H. Steinar, E. Kline, and T. Refice. Evaluating IPv6 adoption in the Internet. In *Proc. Passive and Active Measurement Workshop (PAM)*, 2010.

[19] L. Colitti, S. H. Gunderson, E. Kline, and T. Refice. Evaluating IPv6 adoption in the Internet. In *Proc. PAM*, 2010.

[20] J. Czyz, M. Allman, S. Iekel-Johnson, E. Osterweil, and M. Bailey. Assessing IPv6 adoption. Tec. Report TR-13-004, International Computer Science Institute, 2013.

[21] J. Czyz, M. Allman, J. Zhang, S. Iekel-Johnson, E. Osterweil, and M. Bailey. Measuring ipv6 adoption. In *Proc. of ACM SIGCOMM*, 2014.

[22] Czyz, Jakub and Allman, Mark and Zhang, Jing and Iekel-Johnson, Scott and Osterweil, Eric and Bailey, Michael. Measuring IPv6 Adoption. Technical report, ICSI TR-13-004, 2013.

[23] A. Dainotti, K. Benson, A. King, kc claffy, M. Kallitsis, E. Glatz, and X. Dimitropoulos. Estimating Internet address space usage through passive measurements. *ACM SIGCOMM Comp. Comm. Rev.*, 44(1):43–49, 2014.

[24] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) Specification. RFC 1883 (Proposed Standard), Dec. 1995. Obsoleted by RFC 2460.

[25] A. Dhamdhere, M. Luckie, B. Huffaker, k. claffy, A. Elmokashfi, and E. Aben. Measuring the deployment of IPv6: topology, routing and performance. In *Proc. ACM IMC*, 2012.

[26] Dhamdhere, Amogh and Luckie, Matthew and Huffaker, Bradley and Elmokashfi, Ahmed and Aben, Emile and others. Measuring the deployment of IPv6: topology, routing and performance. In *Proc. ACM conference on Internet measurement conference (IMC)*, 2012.

[27] C. Donley, Ed., L. Howard, V. Kuarsingh, J. Berg, and J. Doshi. Assessing the impact of carrier-grade NAT on network applications. RFC 7021 (Informational), 2013.

[28] A. Dul. Economics of IPv4 transfer market on IPv6 deployment, 2011. (Available at www.quark.net/links.html).

[29] A. Durand, R. Droms, J. Woodyatt, and Y. Lee. Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion. RFC 6333 (Proposed Standard), Aug. 2011.

[30] B. Edelman. Running out of numbers: Scarcity of IP addresses and what to do about it. In *Proc. First Conference on Auctions, Market Mechanisms and Their Applications*, 2009.

[31] H. Elmore, L. J. Camp, and B. Stephens. Diffusion and Adoption of IPv6 in the ARIN Region, 2008. (Available at `http://goo.gl/PgdvKF`).

[32] Google IPv6 statistics. (Available at `http://goo.gl/nl9ZjT`).

[33] Google DNS white-listing. (Available at `http://goo.gl/0heeTs`).

[34] R. Guérin and K. Hosanagar. Fostering IPv6 migration through network quality differentials. *ACM SIGCOMM Computer Communication Review*, 40(3):17–25, 2010.

[35] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister. Census and survey of the visible Internet. In *Proc. ACM IMC*, 2008.

[36] K. F. Hirji, C. R. Mehta, and N. R. Patel. Computing distributions for exact logistic regression. *Journal of the American Statistical Association*, 82(400):1110–1117, 1987.

[37] L. Howard. Internet access pricing in a post-IPv4 runout world. Technical report, 2013. (Available at http://goo.gl/wp0q5i).

[38] G. Huston. IPv6 on OSes. (Available at `http://goo.gl/Upwf78`).

[39] G. Huston. Is the transition to IPv6 a 'market failure'?, 2009. Available at http://www.potaroo.net/ispcol/2009-09/v6trans.html.

[40] S. Iekel-Johnson, C. Labovitz, D. McPherson, and H. Ringberg. Tracking the IPv6 migration. Tech. Report TR-2008-01, Arbor Networks, 2008.

[41] Current Status of IPv6 Support for Networking Applications. (Available at `http://goo.gl/2oThDy`).

[42] Comparison of IPv6 Application Support. (Available at `http://goo.gl/Zl42XJ`).

[43] Comparison of IPv6 support in operating systems. (Available at `http://goo.gl/WTWzhU`).

[44] IPv6 Task Force Report. (Available at `http://goo.gl/yvjqh0`).

[45] S. Jiang, D. Guo, and B. Carpenter. An Incremental Carrier-Grade NAT (CGN) for IPv6 Transition. RFC 6264 (Informational), June 2011.

[46] Juniper IPv6 Solution. (Available at `http://goo.gl/4nYKKZ`).

[47] C. Kalogiros, I. Papafili, G. Stamoulis, C. Courcoubetis, G. Thanos, M. Waldburger, P. Poullie, B. Stiller, D. Field, and M. Bonivace. Final report on economic future Internet coordination activities. Technical Report D2.2-v2.0.doc, 2012. (Available at http://goo.gl/rGsyTD).

[48] E. Karpilovsky, A. Gerber, D. Pei, J. Rexford, and A. Shaikh. Quantifying the extent of IPv6 deployment. In *Proc. PAM*, 2009.

[49] D. Karrenberg. Provider Independent vs Provider Aggregatable Address Space. Technical Report RIPE-127, 1995. (Available at `http://goo.gl/I2Gz3B`).

[50] M. Kuhn and K. Johnson. *Applied predictive modeling*. Springer, 2013.

[51] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet inter-domain traffic. In *Proc. ACM SIGCOMM*, 2010.

[52] K. Larntz. Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics. *Journal of the American Statistical Association*, 73(362):253–263, 1978.

[53] Y. Liu and J. Pan. The impact of NAT on BitTorrent-like P2P systems. In *Proc. IEEE P2P*, 2009.

[54] V. Mahajan and E. Muller. Innovation diffusion and new product growth models in marketing. *Journal of Marketing*, 54(1):55–68, 1990.

[55] C. R. Mehta and N. R. Patel. Exact logistic regression: theory and examples. *Statistics in medicine*, 14(19):2143–2160, 1995.

[56] IPv6 Support in Microsoft Products and Services. (Available at `https://technet.microsoft.com/en-us/network/hh994905.aspx`).

[57] S. Narayan, S. Kolahi, Y. Sunarto, D. Nguyen, and P. Mani. Performance comparison of IPv4 and IPv6 on various Windows operating systems. In *Proc. ICCIT*, 2008.

[58] S. Narayan, P. Shang, and N. Fan. Performance evaluation of IPv4 and IPv6 on Windows Vista and Linux Ubuntu. In *Proc. IEEE NSWCTC*, 2009.

[59] M. Nikkhah and R. Guérin. Migrating the Internet to IPv6: An Exploration of the When and Why.

[60] M. Nikkhah and R. Guérin. Migrating to IPv6 - The Role of Basic Coordination. Tech. report, University of Pennsylvania, 2013. (Available at `http://goo.gl/gWJd95`).

[61] M. Nikkhah and R. Guérin. Migrating to IPv6 - The Role of Basic Coordination. In *Proc. IFIP Networking Conference*, 2014.

[62] M. Nikkhah and R. Guérin. Migrating to IPv6-The Role of Basic Coordination. In *Proc. IFIP Networking*, 2014.

[63] M. Nikkhah, R. Guérin, Y. Lee, and R. Woundy. Assessing IPv6 through web access a measurement study and its findings. In *Proc. ACM CoNEXT*, 2011.

[64] Internet addressing: Measuring deployment of IPv6. Tech. report, OECD, 2010. (Available at www.oecd.org/sti/ict/ipv6.

[65] A. Ozment and S. Schechter. Bootstrapping the adoption of Internet security protocols. In *Proc. WEIS*, Cambridge, UK, June 2006.

[66] P. Peduzzi, J. Concato, E. Kemper, T. R. Holford, and A. R. Feinstein. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, 49(12):1373–1379, 1996.

[67] M. Rappa. Business models on the web, 2000. (Available at `http://goo.gl/iTCWcY`).

[68] IPv6 enabled networks. (Available at `http://v6asns.ripe.net/v/6`).

[69] J.-C. Rochet and J. Tirole. Platform competition in two-sided markets. *J. European Econ. Assoc.*, 1(4), 2003.

[70] J.-C. Rochet and J. Tirole. Two-sided markets: A progress report. *The RAND J. Econ.*, 37(3):645–667, 2006.

[71] R. Roson. Two-sided markets: A tentative survey. *Review of Network Economics*, 4(2), 2005.

[72] N. Sarrar, G. Maier, B. Ager, R. Sommer, and S. Uhlig. Investigating IPv6 traffic – what happened at the World IPv6 Day? In *Proc. PAM*, 2012.

[73] E. Schurman and J. Brutlag. Performance related changes and their user impact, 2009. Available at http://velocityconf.com/velocity2009/public/schedule/detail/8523.

[74] R. L. Smith. A survey of nonregular problems. *Bull. Internat. Statist. Inst*, 53:353–372, 1989.

[75] Site speed: case studies, tips and tools for improving your conversion rate. (Available at http://goo.gl/XV3dmI).

[76] D. Thaler and B. Aboba. What Makes For a Successful Protocol? RFC 5218 (Informational), July 2008.

[77] T-Mobile Goes IPv6 Only on Android 4.4 Devices. (Available at http://goo.gl/WZUBB6).

[78] T. Trinh, L. Gyarmati, and G. Sallai. Migrating to IPv6: A game-theoretic perspective. In *Proc. IEEE LCN*, 2010.

[79] T. A. Trinh, L. Gyarmati, and G. Sallai. Migrating to IPv6: A game-theoretic perspective. In *Proc. IEEE LCN*, 2010.

[80] I. van Beijnum. The future is forever: the state of IPv6 in the Apple world. *ars technica*, May 2012.

[81] T. van der Ploeg, P. C. Austin, and E. W. Steyerberg. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC medical research methodology*, 14(1):137, 2014.

[82] IPv6 at Verizon Wireless. (Available at http://goo.gl/t3hd9i).

[83] World IPv6 Launch Measurements. (Available at http://goo.gl/l1RJil).

[84] E. Vittinghoff and C. E. McCulloch. Relaxing the rule of ten events per variable in logistic and cox regression. *American journal of epidemiology*, 165(6):710–718, 2007.

[85] Configuring IPv6 and IPsec on vSphere ESX, ESXi 4.1 and ESXi 5.x (1021769). (Available at `kb.vmware.com/kb/1021769`).

[86] Y. Wang, S. Ye, and X. Li. Understanding current IPv6 performance: a measurement study. In *Proc. IEEE ISCC*, 2005.

[87] Y. Wang, S. Ye, and X. Li. Understanding current IPv6 performance: A measurement study. In *Proc. 10th IEEE Symp. Comp. Comm. (ISCC)*, 2005.

[88] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62, 1938.

[89] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

[90] J. Wu, J. H. Wang, and J. Yang. CNGI-CERNET2: an IPv6 deployment in China. *ACM SIGCOMM Computer Communication Review*, 41(2):48–52, 2011.

[91] S. Zeadally and L. Raicu. Evaluating IPv6 on Windows and Solaris. *Internet Computing, IEEE*, 7(3):51–57, 2003.

[92] O. Zeitouni, J. Ziv, and N. Merhav. When is the generalized likelihood ratio test optimal? *Information Theory, IEEE Transactions on*, 38(5):1597–1602, 1992.

[93] X. Zhou, M. Jacobsson, H. Uijterwaal, and P. Van Mieghem. IPv6 delay and loss performance evolution. *Intl. J. Comm. Sys.*, 21(6):643–663, 2008.

[94] X. Zhou and P. Van Mieghem. Hopcount and E2E delay: IPv6 versus IPv4. In *Proc. PAM*, 2005.